The Pennsylvania State University

The Graduate School

College of Earth and Mineral Sciences

STREAM CONDITION AND NUTRIENT RUNOFF: LINKING THE SOIL AND WATER ASSESSMENT TOOL (SWAT) MODEL WITH EMPIRICAL ECOLOGICAL MEASURES IN AN AGRICULTURAL WATERSHED IN CENTRAL PENNSYLVANIA

A Thesis in

Geography

by

Claire C. Hirt

© 2016 Claire C. Hirt

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2016

The thesis of Claire C. Hirt was reviewed and approved* by the following:

Robert P. Brooks
Professor of Geography and Ecology
Thesis Adviser

Joseph A. Bishop
Research Associate of Geography

Tameria L. Veith
Agricultural Engineer of United States Department of Agriculture

Cynthia A. Brewer
Professor of Geography
Head of the Department of Geography

*Signatures are on file in the Graduate School.

ABSTRACT

The management of nonpoint sources of nutrients is the primary challenge for improving conditions within the Susquehanna-Chesapeake Basin. Although ecological indicators such as aquatic macroinvertebrates are widely used to assess stream integrity, the relationship between nutrient runoff and stream ecological response remains indistinct. This ambiguity is due, in part, to the logistical and financial hurdles of obtaining high-resolution empirical measurements of nutrients to compare to ecological metrics. However, landscape models that simulate runoff may be a more practical alternative. The Soil and Water Assessment Tool (SWAT) is one such model and was used in this study to simulate nutrient runoff in a small, upland, agricultural watershed in east-central Pennsylvania that drains into Mahantango Creek, a tributary to the Susquehanna River. Macroinvertebrate communities were sampled at 13 sites during May 2014 and described using an Index of Biotic Integrity (IBI). A broad range of scores was observed, indicating differential levels of degradation within the study area that ranged from low to very high. The distribution of IBI scores also indicated that biological integrity was moderately degraded in many reaches throughout the study area. Land cover metrics and the Stream-Wetland-Riparian (SWR) Index, an in-field rapid assessment of site condition, both indicated that agricultural stressors were a major cause of this degradation.

Thirty spatio-temporal scales of nutrient data were identified from SWAT model output, and three nutrients were considered in the analyses: total phosphorus, total nitrate, and total organic nitrogen. Best Subsets Regression was performed on the IBI scores using SWAT, land cover, and SWR variables, and regression models were found to be highly significant. Grab samples of nitrate concentrations were also collected once per season at each site and compared with modeled nitrate levels. In contrast to the SWAT-modeled nutrient outputs, grab sample

nitrate showed no significant relationship with IBI score, supporting the premise that sporadic empirical sampling is an insufficient measure of overall nutrient pollution and is inadequate for characterizing the relationship between nutrient runoff and macroinvertebrate community integrity. This study is among the first to illustrate the value of the SWAT model as a viable alternative to empirical sampling of nutrients in aquatic systems. It also demonstrates the importance of considering spatial and temporal scales of modeled nutrient outputs and additional factors, such as land use and site condition, when describing the relationship between nutrients and stream ecological health.

TABLE OF CONTENTS

LIST OF FIGURES vi
LIST OF TABLES vii
ACKNOWLEDGEMENTS viii
INTRODUCTION 1
Chesapeake Bay Overview 1
Nutrient Impacts 5
Ecological Monitoring and Assessment 6
Nutrient Modeling 11
Research Questions 14
METHODS 17
Study Area 17
Ecological Assessments 21
Landscape Level: Percent Land Cover 21
Rapid Level: Stream-Wetland-Riparian Index 21
Intensive Level: Macroinvertebrate Index of Biotic Integrity 23
Additional Empirical Measurements 25
Soil and Water Assessment Tool 26
Variable Reduction 30
Multiple Regression 31
Additional Comparisons 33
RESULTS 34
SWAT Output Scales 34
SWAT Variable Reduction by Spatio-Temporal Scale 38
Ecological Assessments 42
Landscape Assessment 42
Stream-Wetland-Riparian Index Score 46
Macroinvertebrate Index of Biotic Integrity 48
Regression 52
DISCUSSION 58
LITERATURE CITED 72

LIST OF FIGURES

Figure 1. Extent and location of the Chesapeake Bay Watershed 2

Figure 2. Wedge-shaped relationship between number of macroinvertebrate taxa and nitrate concentration from Wang et al. 2007 10

Figure 3. Hydrologic Response Unit (HRU) generation from a unique intersection of land use, soils, and slope class spatial layers 12

Figure 4. Location of the WE38 study watershed in east-central Pennsylvania 17

Figure 5. Locations of 14 ecological sampling points in WE38 20

Figure 6. Sampling sites and transect subbasins delineated in TopoSWAT 29

Figure 7. The six spatial scales of SWAT output considered in this study 36

Figure 8. Percent Forest for each sampling site at six landscape scales 45

Figure 9. Stream-Wetland-Riparian (SWR) Index Score for each sampling site 47

Figure 10. Index of Biotic Integrity (IBI) Score for each sampling site 48

Figure 11. Nutrient level, Percent Forest, SWR Score, and IBI Score for each sampling site..... 65

LIST OF TABLES

Table 1. Metrics used in the Stream-Wetland-Riparian (SWR) Index 22

Table 2. The six component metrics to the PA-DEP Index of Biotic Integrity. 25

Table 3. Spatial and temporal scales of SWAT model output for nutrient runoff. 38

Table 4. The final five representative SWAT variables after Round Two elimination. 39

Table 5. The final five representative SWAT variables and the other SWAT variables with which they are correlated (Spearman's rho, p<0.05) 39

Table 6. The eight representative variables remaining after Round 1 correlation elimination..... 41

Table 7. Spatial scales of analysis for land cover classes 43

Table 8. Regression of Percent Agriculture and Percent Forest within each spatial scale 43

Table 9. Correlation matrix of Percent Forest 46

Table 10. Correlation matrix of overall SWR Score with component metrics 48

Table 11. Correlation matrix of Floodplain-Wetland (FPWL) Score with component metrics ... 48

Table 12. Correlation matrix of IBI Score with component metrics 50

Table 13. Water chemical and physical properties 51

Table 14. Grab sample nitrate concentrations..... 52

Table 15. Correlation matrix of average and seasonal grab sample nitrate concentrations..... 52

Table 16. Best Subsets Regression of the final five SWAT variables on IBI Score 53

Table 17. Best Subsets Regression of SWAT and Percent Forest on IBI Score 54

Table 18. Best Subsets Regression of SWAT, Land Cover, and SWR on IBI Score..... 55

Table 19. Best Subsets Regression of SWAT, Land Cover, SWR, and Grab Sample Nitrate on IBI Score 56

Table 20. Comparison of correlated SWAT variables in final model of IBI..... 57

ACKNOWLEDGEMENTS

This thesis would not have been possible without the guidance, contributions, and support of many people. First and foremost, I would like to thank Dr. Robert Brooks for advising this project from start to finish and for providing this research opportunity. I am grateful for his invaluable advice and unwavering optimistic encouragement every step of the way. Dr. Joseph Bishop and Dr. Tamie Veith have also contributed a great deal as committee members, and I am grateful for all of their constructive feedback. Joe taught me many of the GIS skills required to do the analyses here, and Tamie oversaw the incorporation and interpretation of the SWAT modeling in the project.

Dr. Amy Collick also deserves a special mention, as this research truly would not have been feasible without her dedicated work on the SWAT modeling. I am grateful for the many long days she spent writing and editing code and for her remarkable patience in trying to explain SWAT to a fledgling ecologist. I would also like to thank Ms. Suzy Yetter for generously sharing her vast macroinvertebrate knowledge with me. She has been a mentor in countless ways and her ideas and feedback were fundamental to troubleshooting aspects of this thesis and to interpreting the ecological data. The other members of our research group, Riparia, have also added significantly to this project. Dr. Mike Nassry, Ms. Hannah Ingram, Mr. Andrew Townsend, and Ms. Tara Mazurczyk were a great help in conducting the rapid assessments and sifting through the large SWAT datasets. I owe a special thank you to Ms. Jackie Dougherty, our undergraduate intern who spent her summer working long days in the field and sorting through pans of stream debris with me.

I would also like to thank the folks at the USDA Agricultural Research Service in State College, PA, whose collaboration on this project made for an enriching experience and a notable

example of how rewarding multidisciplinary research can be. Thank you to Dr. Pete Kleinman for offering his feedback on this project, to Dr. Tony Buda for sharing his hydrology expertise, and to Mrs. Sarah Fishel for helping with the grab sample processing. Mr. Todd Strohecker was also a key figure in this project, acquiring sites and familiarizing me with the watershed. I would also like to thank the landowners of WE38 that granted access to their properties and, as a result, have allowed this contribution to freshwater science to come to fruition.

This thesis was part of a larger grant project called the Center for Nutrient Solutions, a research effort to manage nutrient pollution issues in the Chesapeake Bay watershed. I am grateful to the team of investigators that listened attentively to my presentations at meetings and provided productive feedback on how to improve and proceed. I would like to acknowledge the Environmental Protection Agency as the major funding source of this research (Grant Number RD835568), as well as the Penn State University Department of Geography for additional funding. Thank you to the Ruby S. Miller Fund, the C. Gregory Knight Fund, and Dr. Erica Smithwick's Center for Landscape Dynamics for the awards that made travel to conferences to present this research possible.

Lastly I would like to acknowledge my most indispensable supporters – my family and friends. Thank you to my parents for believing in me and encouraging me to pursue my passion for science from the moment I caught my first bug. Thank you to my siblings, in-laws, and friends for their boundless support and care. Most importantly, thank you to my husband, Matt, who was there for me through every twist and turn of graduate school and made completing this thesis possible in innumerable big and small ways. I could not have done it without you.

INTRODUCTION

Chesapeake Bay Overview

Management of nutrient runoff is the foremost challenge to improving aquatic conditions in the Susquehanna River-Chesapeake Bay watershed. The Chesapeake Bay watershed has an area of 164,200 km^2 and contains major population centers, including Washington, D.C. and Baltimore, MD (Figure 1). The Bay itself has changed dramatically due to pervasive and severe anthropogenic impacts. Since pre-colonial times, inputs of nitrogen have increased six- to eight-fold, and inputs of phosphorus have increased 13- to 24-fold (Boynton et al. 1995), especially due to the increase in commercial fertilizer use and import of animal feed since the 1950s (Kemp 2005). This estuary experiences serious eutrophication problems, including anoxic and hypoxic conditions that occur in deepwater and nearshore zones. Aquatic habitat has been reduced and fish kills and nuisance and toxic algal blooms occur annually (Bricker et al. 2007). The primary nutrient sources are agriculture, wastewater, urban runoff, septic tanks, sewer overflows, and atmospheric deposition (Bricker et al. 2007).

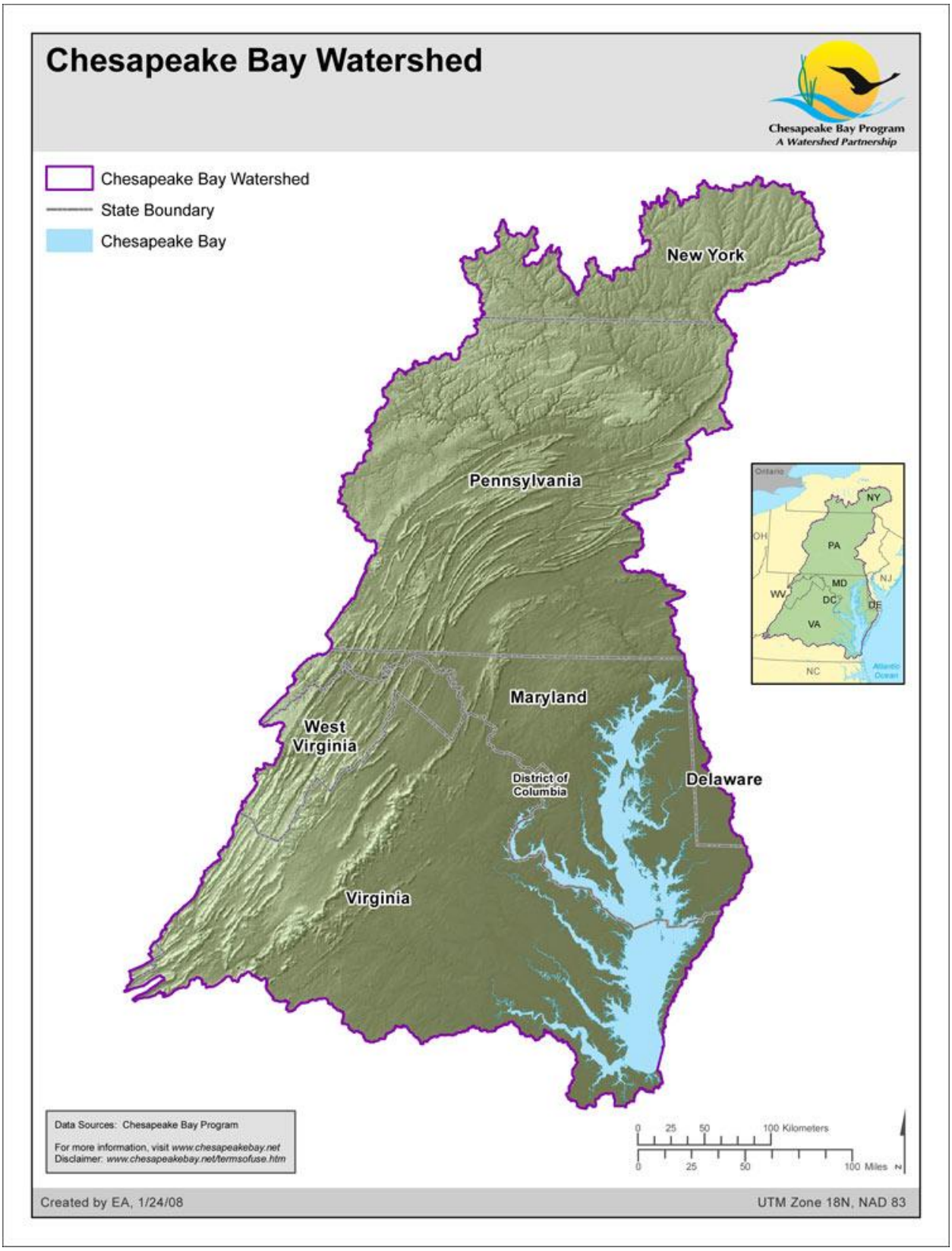


Figure 1. Extent and location of the Chesapeake Bay Watershed. Source: Chesapeake Bay Program, January 2008, Web.

In the U.S., aquatic systems are federally regulated by the Clean Water Act (1972), which requires the maintenance and restoration of the "chemical, physical, and biological integrity of the Nation's waters" (Clean Water Act 1972, SEC. 101. (a)). The Chesapeake Bay was singled out for supplementary legislation with the Chesapeake Bay Agreement of 1983, when it was the first estuary to be targeted for protection and restoration by Congress (CBP 1983). The 1987 Chesapeake Bay Agreement joined state and federal groups in a commitment to develop guidelines and implement programs to improve and protect water quality and habitat conditions. Specifically, the policy aimed to assess and control nonpoint source pollution and to reduce nutrient inputs by 40% by 2000 (CEC 1987). Later, the Chesapeake 2000 Agreement committed to defining the specific conditions protective of living resources, which increased these required reductions to 48% and 53% for nitrogen and phosphorus, respectively (CEC 2000, USEPA 2003).

In the last six years, efforts to improve the Bay's condition have been recharged with President Barack Obama's issuance of Executive Order 13508 (2009). The Order on Chesapeake Bay Protection and Restoration mandated development and implementation of a fresh strategy for the Bay watershed and region. The overarching goals of The Order were to restore clean water, recover habitat, sustain fish and wildlife, and expand citizen stewardship (FLC 2014). Reducing nitrogen and phosphorus are paramount to achieving these objectives. To this end, the Total Maximum Daily Load (TMDL) Accountability Framework was established in 2010 to regulate both point and nonpoint sources across all seven jurisdictions within the Chesapeake Bay watershed (USEPA 2010a, Linker et al. 2013). Watershed Implementation Programs (WIPs) for each of the jurisdictions were developed to reduce loads to allocated TMDL levels by 2025 (USEPA 2010b, Linker et al. 2013).

Progress in improving aquatic condition has been observed recently, and nutrient and sediment loads have decreased in the region. In 2013, conservation practices were established on more than 110,000 ha (271,000 ac) within the Bay watershed, totaling approximately 526,000 ha (1.3 million ac) to date. In 2012, 900 ha (2,231 ac) of wetlands were also established or restored on agricultural lands within the watershed (FLC 2014). Nutrient load reductions have been observed in 72% of tributary rivers for nitrogen and 81% of tributary rivers for phosphorus (Langland et al. 2006). Wastewater treatment plants have been updated, atmospheric nitrogen deposition has decreased, and agricultural nutrient loadings to streams have been reduced (Jaworski 1990, Lyerly et al. 2013). The ecology of the Bay has shown improvement, including some recovery in submerged aquatic vegetation (SAV), a primary indicator for the estuary, in the upper reaches of the Bay (Bricker et al. 2007) and a decrease in the duration of summer hypoxic zones from five to four months (Zhou et al. 2014).

These efforts and results are encouraging, but continued improvements remain essential. To date, most reductions have come from the management of point sources, with far less progress in managing nonpoint sources (Boesch et al. 2001). This is a concern because two-thirds of phosphorus and three-fourths of nitrogen in the Bay comes from nonpoint sources (CENR 2010). Activities in agricultural and urban areas contribute to nutrient levels (Carpenter et al. 1998), but agriculture is the principal source of nonpoint nutrient pollution in the region (USEPA SAB 2011). Fertilizer use, nitrogen-fixing crops, and animal feed imports are major drivers (Boyer et al. 2002, CENR 2010). Nutrient discharge is positively correlated with proportion of cropland (Jordan et al. 1997, Chambers et al. 2012) and animal density (Carpenter et al. 1998). Nonpoint sources of nutrients have far-reaching impacts, including contributions to the large-scale hypoxia in the Bay (CENR 2010).

Between 2010 and 2012, only 29% of the Chesapeake Bay estuary had attained water quality standards for dissolved oxygen, water clarity/underwater bay grasses and chlorophyll-a. The goal is to reach 60% attainment by 2025 (FLC 2014). It is clear that upgrades to wastewater treatment plants and improvements to other point sources will not be enough with nonpoint activities as dominant sources of nutrients (Lyerly et al. 2013).

Nutrient Impacts

Worldwide, humans have approximately doubled the rate of nitrogen input into the terrestrial nitrogen cycle through fertilizer synthesis and application, leguminous crop cultivation, and fossil fuel burning (Vitousek et al. 1997). This bioavailable nitrogen continues to increase (Vitousek et al. 1997, Mulholland et al. 2008). Excess nitrogen is highly mobile and can leach into aquatic systems or volatize in the atmosphere to be deposited elsewhere. Excess phosphorus accumulates in the soil, and a portion is transported to aquatic systems, mostly via surface flow adsorbed to sediment (Jordan et al. 1997, Carpenter et al. 1998, Correll 1998).

The overall impacts of nutrient excesses include increased global greenhouse gas concentrations from nitrogen dioxide (NO_2), leaching of other essential soil nutrients, loss of biodiversity, and acidification of soils, streams, and lakes (Vitousek et al. 1997, USEPA SAB 2011). In aquatic systems, eutrophication is associated with increased algal abundance, loss of sensitive benthic macroinvertebrate taxa, and increased benthic diatom taxa indicative of eutrophication (Chambers et al. 2012). Algal blooms can be toxic and have cascading ecological effects. As blooms occur, benthic sunlight diminishes, submerged aquatic vegetation (SAV) declines, and habitat and resources for higher trophic levels decrease. As algae die, sink, and decay, the benthic dissolved oxygen decreases, causing hypoxic or anoxic conditions (CENR

2010). Since the 1950s hypoxia has increased substantially, with an order of magnitude increase in the number of water bodies worldwide affected by low oxygen (Boesch et al. 2001, Kemp 2005, Diaz and Rosenberg 2008).

Drinking water quality is regulated for nitrate because at high concentrations it can restrict oxygen transport in the bloodstream of young infants, a condition called methemoglobinemia or, more commonly, "blue baby syndrome" (Knobeloch et al. 2000, USEPA SAB 2011). Shallow aquifers pose a particular risk due to limited natural filtration through shallow soils, and a portion of tested wells have been found to exceed the regulatory threshold (Dubrovsky and Hamilton 2010). Additionally, algal blooms can cause taste and odor problems or even respiratory distress and neurological problems in swimmers (USEPA 2007).

Ecological Monitoring and Assessment

The Clean Water Act (1972) placed legal protection on the integrity of three important aspects of aquatic systems: the chemical, physical, and biological. The assessment of physical and chemical water parameters tends to be relatively straightforward to describe, measure, and begin to address. However, biological integrity was and continues to be a more tenuous target and is perhaps the most important of the three to the long-term health of the Bay, or any aquatic ecosystem. Physical condition alone can overlook stressors such as acid deposition, which has severe effects on biota but leaves the abiotic environment relatively unaffected. Chemical condition alone is both impractical and insufficient; every chemical cannot be evaluated, lotic water chemistry is transient, synergistic effects on biota occur, and frequent monitoring is expensive and logistically difficult (PA-DEP 2009). Biological integrity refers to an ecosystem's capacity to sustain its full range of elements and processes that would be expected in areas with

minimal human influence (Karr and Dudley 1981, PA-DEP 2009). Protecting biological integrity in turn preserves human uses (Karr and Chu 2000). Assessment of biological integrity must include monitoring of biological condition explicitly.

The index of biotic integrity (IBI) was developed by Karr (1981) to evaluate biological integrity in stream ecosystems. Combining multiple metrics that characterize the biological community within a stream reach, scores are compared to regional reference standards to quantify impacts. Over the past three decades, the IBI has become a widely used tool for assessing aquatic system condition, especially for its ability to guide management decisions. For example, in Pennsylvania, IBI scores help to designate impaired stream segments that require a TMDL to decrease pollution and return water to its full use and function (PA-DEP 2012a).

While Karr's first IBI was focused on fish, IBIs have since been developed for many taxonomic groups and numerous ecosystem types. The most frequently used ecological group for monitoring stream ecological condition today is aquatic benthic macroinvertebrates (hereafter macroinvertebrates) (Rosenberg and Resh 1993). Macroinvertebrates are aquatic organisms that lack internal bone structure, are macroscopically visible, and reside on the bottom substrate. They play a vital role in transferring energy from lower to higher trophic levels. They are widely used for IBIs because they are ubiquitous, highly diverse, long-lived, sedentary, and exhibit differential tolerances and sensitivities to stressors (Rosenberg and Resh 1993, Barbour et al. 1999, PA-DEP 2009). Examining the composition of the macroinvertebrate community can elucidate the pollution level of that system. If pollution sensitive taxa are lost, and pollution tolerant taxa are gained, this indicates that pollutants are stressing the system (Karr 1981, Barbour et al. 1999, PA-DEP 2009). Monitoring macroinvertebrate communities in stream ecosystems reflects the overall condition from multiple stressors over time (Karr 1991).

Many regulatory agencies throughout the United States and abroad have active and ongoing monitoring programs to assess biotic condition using benthic macroinvertebrates, including the jurisdictions of the Chesapeake Bay Watershed. In Pennsylvania, it is the Department of Environmental Protection Statewide Surface Water Quality Monitoring Program (PA-DEP 2012b). Maryland has the Maryland Biological Stream Survey within the Department of Natural Resources (Klauda et al. 1998). In Virginia, the Department of Environmental Quality runs the Freshwater Biological Monitoring Program (Burton and Gerritsen 2003). New York has the Stream Biomonitoring Unit of the Department of Environmental Conservation (NYS-DEC 2012). And, in West Virginia, the Department of Environmental Protection Watershed Assessment Branch conducts their monitoring and assessment using macroinvertebrates (WV-DEP 2015).

These regulatory agencies use macroinvertebrate sampling to calculate IBI scores that designate overall aquatic condition. Scores can then be tied to impairment. In Pennsylvania, for example, PA-DEP has determined that an IBI score of <50 automatically designates stream impairment (PA-DEP 2012a). Impaired streams are then placed on their 303(d) list as required by the Clean Water Act and the reason for impairment is also listed. The impairments may be point or nonpoint sources, including agricultural runoff of nitrogen or phosphorus. States are then required to develop TMDL plans to address impaired waters on the 303(d) list and submit them to the U.S. Environmental Protection Agency (EPA) for approval. The goal of the TMDL is to place a limit on the maximum quantity of each pollutant that the impaired waterbody can receive to allow it to recover to an unimpaired condition. This total maximum pollutant load is then allocated to the various pollutant sources so that each contributing source is accounted for and limited in its daily pollutant discharge (USEPA 1999).

Impairments from nutrients are obvious when the impacts are significant, such as widespread algal overgrowth on stream substrate. However, when impacts are less physically conspicuous, nutrient pollution can be more difficult to diagnose. Although the most egregious nutrient pollution should certainly be addressed, other moderately nutrient-enriched streams are also in need of nutrient impairment designation and TMDL guidelines to continue making progress on improving aquatic conditions in the Chesapeake Bay and in its smaller contributing watersheds. Because statewide assessments widely use macroinvertebrates to assess ecological condition and designate impairment, it would be beneficial to delineate the link between nutrient enrichment and IBI score. Although a large amount of research has been done on this topic, the conclusions have varied and, at times, conflicted in regards to the importance of nutrients in shaping macroinvertebrate communities (e.g., Wang 2007, Calhoun et al. 2008).

One of the challenges of identifying impacts of nutrients on macroinvertebrates is that nutrients act in conjunction with other stressors (Wang 2007). Habitat quality, percent agriculture, fine sediment, buffer land use, riparian quality, substrate disturbance, livestock access, hydrologic disturbance, and hydromorphology have also been noted as important in shaping macroinvertebrate communities (Richards et al. 1993, Miltner and Rankin 1998, Genito et al. 2002, Riseng et al. 2004, Yuan and Norton 2004, Hering et al. 2006, Kyriakeas and Watzin 2006, Tran et al. 2010, Wagenhoff et al. 2012). The majority of research has shown that, despite these additional confounding factors, macroinvertebrates do exhibit decreasing trends with increasing concentrations of phosphorus and nitrogen (Miltner and Rankin 1998, Riseng et al. 2004, Wang 2007, Justus et al. 2010). A wedge-shaped response is typical, with metrics being highly variable at low nutrient concentrations and consistently impacted at high nutrient concentrations (Figure 2) (Wang 2007, Ashton et al. 2014).

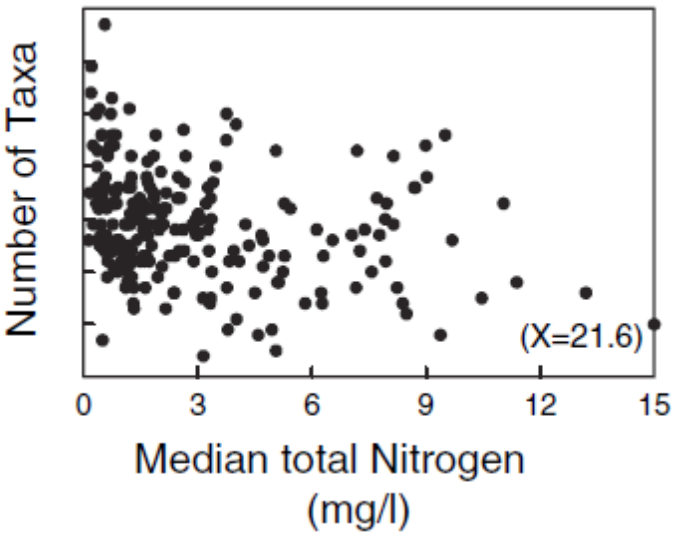


Figure 2. Wedge-shaped relationship between number of macroinvertebrate taxa and nitrate concentration from Wang et al. 2007.

In addition to the presence of confounding factors, nutrient source areas are variable through both space and time, a characteristic that poses significant challenges in obtaining the high-resolution nutrient data necessary to establish a clear link between ecology and nutrients. The financial costs of water chemistry analysis and field technician time are often prohibitively expensive. Researchers may also face intractable opposition to gaining permission for sampling water chemistry on private property, especially in agricultural areas where regulation can be contentious. These roadblocks leave researchers with sporadic and sparse grab sampling as the only option for measuring nutrient pollution. Unfortunately, such sampling approaches are inadequate to capture the fluctuating nature of runoff that may be important in defining the relationship between nutrients and macroinvertebrate communities and may help to explain the noisy results that previous studies have produced (e.g., Richards et al. 1993, Genito et al. 2002).

Nutrient Modeling

Although long-term empirical sampling of nutrient concentrations is not typically feasible, there may be an alternative: landscape modeling. Models have been developed to simulate the processing of nutrients and water across a watershed, producing spatially explicit representations of the origin and amount of runoff within that watershed. The Soil Water Assessment Tool (SWAT) (Neitsch et al. 2002) is one such model. SWAT works within a Geographic Information System (GIS) and incorporates weather, soil, land cover, and land management data to simulate surface and subsurface hydrology along with fluxes of chemicals and sediments (Collick et al. 2014).

The following is an explanation of SWAT from Amy Collick, Ph.D, a modeler with the United States Department of Agriculture Agricultural Research Service (USDA-ARS) who conducted the modeling portion of this study (Personal Communication, July 7, 2015):

> *The Soil and Water Assessment Tool (SWAT) is a physically based watershed model extensively used as a tool to simulate hydrology and nutrient transport in watersheds in the U.S. and abroad. Using spatial datasets of elevation, land use, and soils as well as weather, field management, and the array of default and user-defined watershed parameters contained in SWAT's extensive databases, SWAT simulates the hydrological and nutrient transport processes at a hydrologic response unit (HRU) level. After delineating the watershed, HRUs are generated from a unique intersection of the spatial datasets of land use, soil, and slope classes for each subbasin (Figure 3). SWAT does not route water and*

nutrients from one HRU to another but rather directly into to the subbasin's stream reach to then be routed to the main watershed outlet.

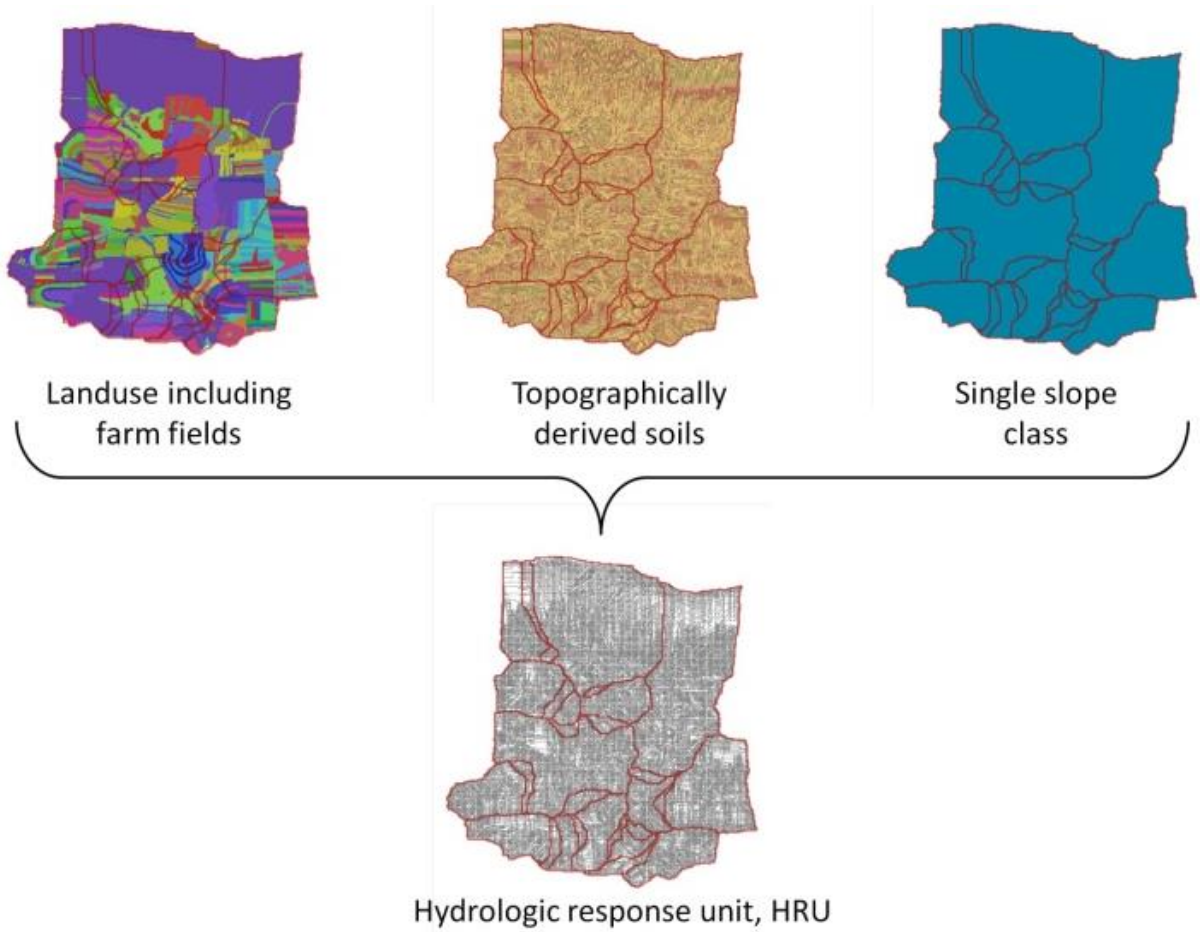


Figure 3. Hydrologic Response Unit (HRU) generation from a unique intersection of land use, soils, and slope class spatial layers for each subbasin of the watershed. The HRU serves as the smallest spatial processing unit within SWAT (Amy Collick, Personal Communication, July 7, 2015)

An ArcGIS interface, aptly called ArcSWAT (http://swat.tamu.edu/software/arcswat/), allows users to integrate the spatial datasets and layers of topography, land use, and soil into the initialization of the model including watershed delineation and HRU generation. A modified version of SWAT has been developed for ArcSWAT called TopoSWAT. It uses an ArcMap

Toolbox (ESRI 2012. ArcGIS Desktop: Release 10.1 Redlands, CA: Environmental Systems Research Institute) to generate a spatial layer of topographic wetness index (TI), which is intersected with the soils layer, to replace the standard SWAT soils and slope class layers. Spatially, TI values can be calculated and mapped from the slope, flow direction, and flow accumulation spatial datasets generated from an elevation dataset (Digital Elevation Model, DEM) for an area or watershed. The TIs are then categorized into ten classes of equal area to represent the soil moisture gradient (1 for the driest to 10 for the wettest) across the watershed. In the resulting TI/soil spatial layer, the soil characteristics, such as bulk density, available water content, depth of horizons, texture, and hydraulic conductivity, are varied across the topographic wetness index classes.

Described initially in Beven and Kirkby (1979) to model variable contributing areas, the topographic wetness index (TI) indicates a landscape area's potential for saturation and consequent influence on runoff generation and is a function of the contributing area, α, of a point and the local slope, β.

$$TI = ln\left(\frac{\alpha}{tan\,\beta}\right)$$

Easton et al. (2008) incorporated TI into ArcSWAT in order to model variable source hydrology thus initiating the development of TopoSWAT. TopoSWAT and the topographically derived soils have been shown to improve hydrology simulations and phosphorus representation in the WE38 watershed in Pennsylvania (Collick et al., 2014) and hydrology in several small subbasins (fields) in an experimental watershed in Riesel, Texas (Fuka et al., 2015).

The use of the SWAT model to link nutrients with biological integrity is a relatively new research direction and holds potential for the fields of aquatic conservation and ecology. Kautza and Sullivan (2012) described the relationship between fish IBI scores and SWAT nutrient outputs in Ohio, indicating that SWAT is a useful tool to complement field-based surveys. Woznicki et al. (2015) explored this relationship in Michigan with fish and macroinvertebrate community metrics. Their models also identified important links between ecological indicators and SWAT-modeled nutrient data. Additional studies in the Michigan region have been conducted by this same research group, all with similar conclusions (see Einheuser et al. 2012, Einheuser et al. 2013a, Einheuser et al. 2013b, Daggupati et al. 2015, Herman et al. 2015, Woznicki et al. 2016). One piece that may be missing from all of these studies, however, is a consideration for the scale at which SWAT nutrient data is summarized. The SWAT model is typically run on a daily timestep over many years for each Hydrologic Response Unit (HRU), which generates an immense number of data points for each nutrient (e.g., Total Organic Nitrogen) at each HRU for each day of the model run. SWAT modelers often provide an average value at one spatial scale, the Reach scale (see SWAT Output Scales, page 34), for each macroinvertebrate sampling point to ecologists, who then examine the relationship between ecological metrics and these averaged values. In this averaging process, temporal and spatial detail that may be influential in shaping macroinvertebrate communities may be lost.

Research Questions

The goal of this research was to take an in-depth look at the SWAT model from an ecological perspective and to test how SWAT-modeled nutrient runoff relates to

macroinvertebrate community integrity, as an endpoint proxy for ecological condition. Specifically, this project was guided by four research questions:

1. How can SWAT model output be summarized and post-processed to provide an array of nutrient runoff data variables at various spatial and temporal scales that may be valuable to ecologists?
2. How related or unique are each of these nutrient variables? Do the various spatial and temporal scales each convey different information about nutrient runoff in WE38, or are some redundant?
3. What is the relationship between SWAT-modeled nutrients and stream ecological condition, as measured by the macroinvertebrate IBI?
4. Because macroinvertebrates are known to be impacted by other variables in addition to nutrients, can the relationship between nutrient runoff and stream condition be improved by including additional land use, site condition, and water chemistry variables? Which of these variables are most relevant to explaining macroinvertebrate community integrity?

The first question, "How can SWAT output be summarized or post-processed?" was approached by examining the SWAT model outputs and meeting with SWAT modelers at the Pasture Systems and Watershed Management Research Unit of the USDA-ARS in State College, PA. The data generated directly from SWAT was considered, as well as the other ways that ecologists can transform these data to produce additional spatio-temporal scales in post-processing of the output. The second question, "How are these nutrient variables related?" was asked because there is presumably an overlap in the information that these spatial and temporal scales convey, since they all originate from the same daily-timestep SWAT data.

The final questions considered the ecological components of this project. Macroinvertebrate communities were assessed using an IBI, and then IBI scores were used to test the relationship between macroinvertebrates and SWAT-modeled nutrients (question three). Additional assessments were also completed to see if this relationship was strengthened in light of other factors beyond modeled nutrient concentrations (question four). These other ecological assessments included landscape-scale land cover metrics, a rapid assessment of stream, wetland, and riparian condition, and water chemistry sampling.

METHODS

Study Area

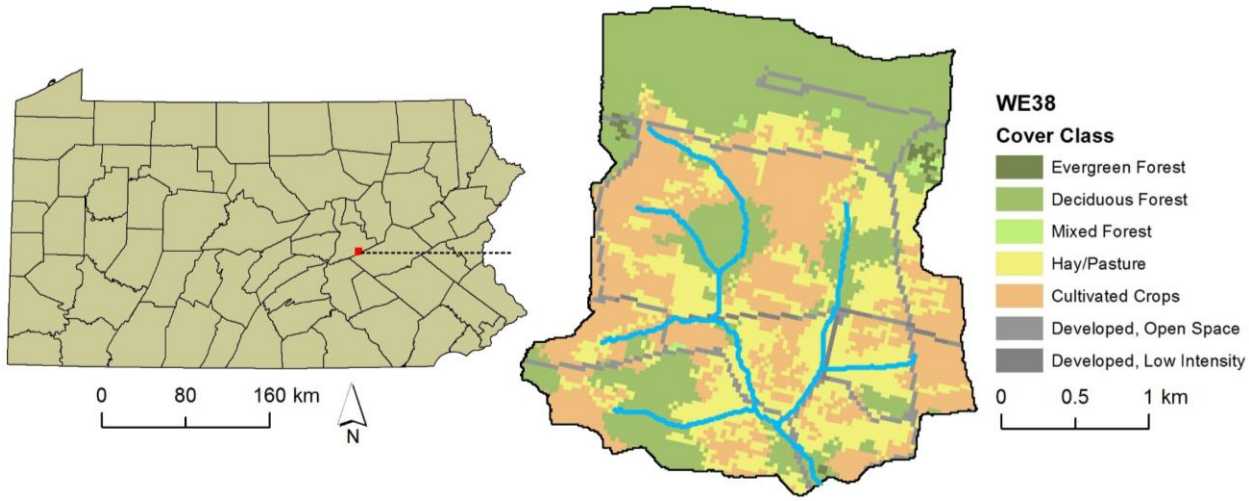


Figure 4. Location of the WE38 study watershed in east-central Pennsylvania. Shown are land cover classes from the 2011 National Land Cover Database (Homer et al. 2015) and stream locations.

This study took place in a small (7.3 km^2) upland agricultural watershed in east-central Pennsylvania labeled WE38 (Figure 4). WE38 was established in 1967 as a USDA-ARS experimental watershed. The area was selected for its Ridge and Valley upland agricultural characteristics and lack of karst geology (USDA-ARS 2011). This watershed is part of the larger (420 km^2) Mahantango Creek watershed that drains into the Susquehanna River, the largest tributary to the Chesapeake Bay. WE38 was an ideal location to assess correlations between SWAT-modeled nutrients and macroinvertebrates for three reasons. First, as a long-term study watershed of the USDA-ARS, there was a wealth of data on the area that translated to higher confidence in model outputs and an overall availability of additional resources. Second, the majority of the watershed area is within headwater regions, a spatial scale relevant to many agricultural management decisions. Third, because the Susquehanna River provides over half of

the freshwater to the Chesapeake Bay (USGS 1999), management of the Susquehanna Basin holds particular importance.

The mean elevation in WE38 is 286 m above sea level, ranging from 215 m in the valleys to 503 m on ridges. The climate is temperate and humid. Annual rainfall averages 1065 mm/yr and stream flow averages 450 mm/yr. Soils include Calvine, Klinesville, Berks, Meckesville, Harleton, Leck Kill, and Albrights. Land use is dominated by agriculture (44.5%), which consists of a mixed cropping system of soybean, small grain, rye, corn, and hay. The farms are privately owned and operated. Forest (38.8%) is also prevalent and is mostly distributed in the northern uplands with some woodlots intermixed between agricultural fields. The remaining land consists of developed areas (6.2%), continuous pasture (3.5%), Conservation Reserve Enhancement Program (CREP) areas (3.1%), Christmas tree plantations (2.2%), and fallow and grasslands (0.9%). Open water comprises less than 1% of watershed area (Collick et al. 2014).

Because the long-term USDA-ARS research program exists, agricultural records, measured data, and previous studies on hydrology and nutrient runoff are abundant for WE38. Detailed agricultural management schedules for fields were provided by farmers for a 12-year period from 1999 through 2010. Extensive hydrological, weather, soil, water quality, and farm management data have also been collected (Bryant 2011). The SWAT model has previously been run on WE38 by modelers at the USDA-ARS Pasture Systems and Watershed Management Research Unit (Collick et al. 2014). The dynamics of nutrient runoff have also been studied using 30 years of experimental and monitoring data and found that 90% of algal-available phosphorus export occurred during the seven largest storms and typically from areas where high soil phosphorus and high erosion intersected near streams. Nitrate was found to be exported in

late winter and spring, after the summer-deficit groundwater recharge was reached (Pionke et al. 2000).

A study on the macroinvertebrate community of WE38 was also conducted previously from 1998 to 2000 by Genito et al. (2002). The objective was to test if macroinvertebrate indicators agreed with physiochemical water quality descriptions. Land use percentages, habitat quality, and nitrate concentrations were used in the multivariate comparisons. Results showed that the percentage of agricultural land and habitat quality were more influential than nitrate in describing the environmental gradient to which macroinvertebrates responded. It is important to note, however, that nitrate concentrations were measured with one annual grab sample at each site. Additionally, phosphorous was not considered because all grab samples were collected during baseflow conditions. My thesis project included higher resolution nutrient data on both nitrogen and phosphorus runoff and thus had the potential to better describe the contributions of nutrients to ecological degradation in the watershed.

To assess the relationship between SWAT outputs and ecology, sampling points were placed upstream of every confluence in the WE38 watershed. Additionally, because the riparian zone of most streams in WE38 are agricultural, stream reaches that were located within forested buffers and patches were also included as sampling points to broaden the variety of sampled habitats. All sampling locations were a minimum of 50 m upstream of confluences or structures (such as culverts or bridges) to minimize effects of these factors on stream velocity, depth, and habitat quality (Barbour et al. 1999). In all, 14 points were selected for ecological sampling (Figure 5).

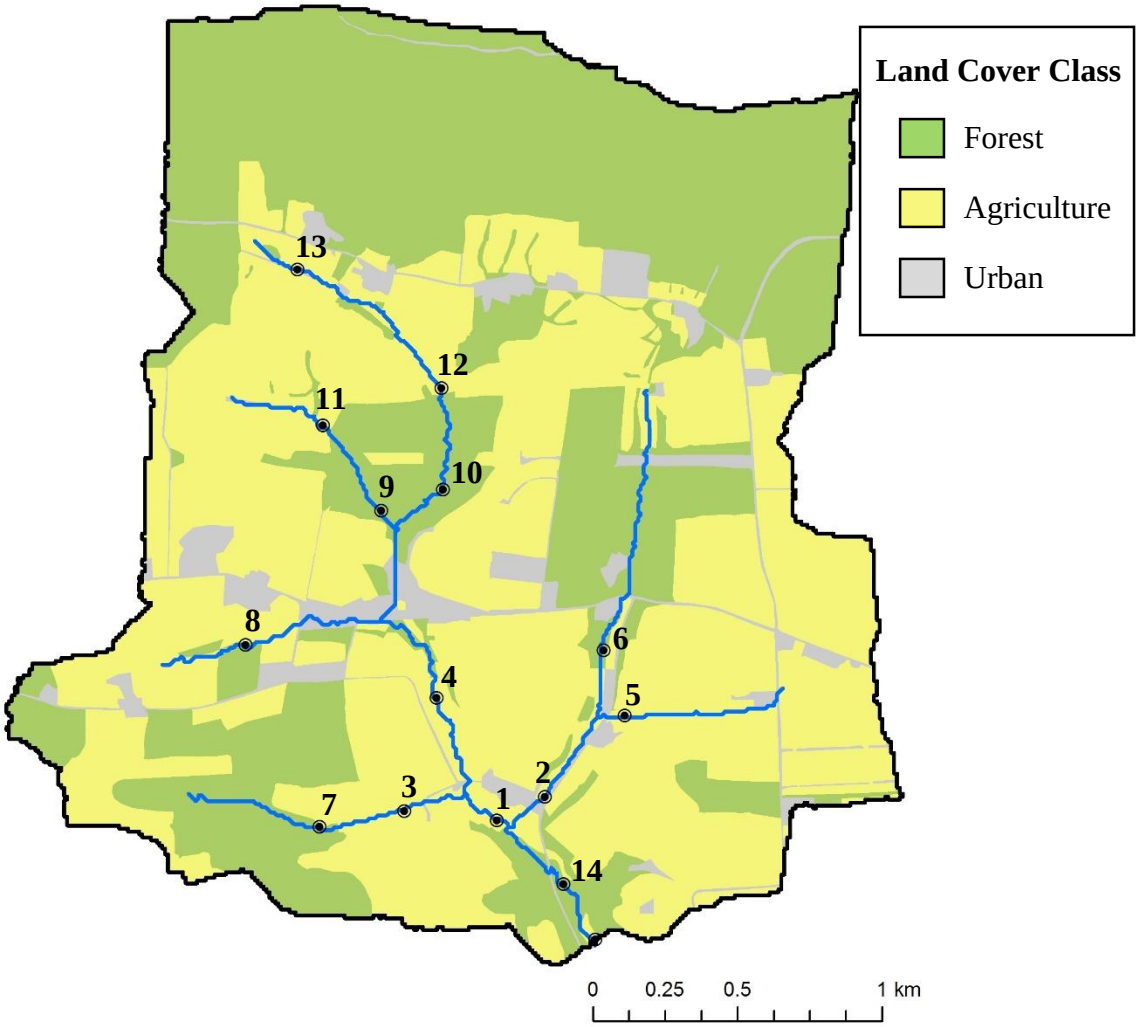


Figure 5. Locations of 14 ecological sampling points in WE38. Streams are shown in blue. Land cover classes are shown in green (forest), yellow (agriculture), and grey (urban).

Ecological Assessments

Three levels of ecological assessments were conducted during this study: Level 1-Landscape, Level 2-Rapid, and Level 3-Intensive (Brooks et al. 2006). Landscape-level assessments can be done quickly, often at broad spatial scales and digitally from the office. The landscape assessment selected was Percent Land Cover. Rapid assessments can be completed in a few hours or less and are typically done in the field. For this study the Stream-Wetland-Riparian Index was used (Brooks et al. 2009). Intensive assessments require extensive time investment and may include both field and laboratory components. The PA-DEP Macroinvertebrate Index of Biotic Integrity for Wadeable, Freestone, Riffle-Run Streams was selected for this study (PA-DEP 2012a).

Landscape Level: Percent Land Cover

The percentages of agriculture, forest, and urban land cover were calculated for each sampling site. A number of spatial scales were included to match the spatial scales identified for SWAT model output, as well as two "landscape circle" scales. These circles were created by generating a 100-m or 1-km radius circle around the center of each sample reach in ArcGIS using the Buffer tool (ESRI 2012). The 2011 National Land Cover Database Land Cover layer was used for all spatial scales (Homer et al. 2015), and percentages were calculated in a geographic information system (GIS) in ArcGIS. See "Results" section for the final list of Landscape Level scales.

Rapid Level: Stream-Wetland-Riparian Index

The Stream-Wetland-Riparian (SWR) Index is a field assessment that can be used alone or in conjunction with intensive benthic macroinvertebrate sampling. SWR is a rapid assessment

of overall aquatic ecosystem condition developed by Brooks et al. (2009) in the Mid-Atlantic Region. Various measurements (Table 1) were taken at field locations to calculate seven component metrics that contribute to overall SWR index score. The average sampling time was between one and two hours.

Table 1. Metrics used in the Stream-Wetland-Riparian (SWR) Index, adapted from Brooks et al. (2009).

Metric	Location	Description
Buffer Condition 0-300 m from stream	Wetland-Riparian	Adjacent land use type
Incision Ratio	Stream	Bankfull parameters with stream channel measurements
Invasives Cover Class	Wetland-Riparian	By species
Basal Area	Wetland-Riparian	3 Bitterlich Tree Points with diameter at breast height measurements for each tree
Number of Wetland-Riparian Stressors	Wetland-Riparian	Checklist of stressors, including location and distance from stream
Number of Stream Stressors	Stream	Checklist of stressors, including location and distance from stream
Stream Habitat Assessment Score	Stream	Based on USEPA Rapid Bioassessment Protocol

Each SWR assessment began by measuring out to a 100-m by 100-m sampling area, centered over the stream channel. Within the wetland and riparian portion of the sampling area, the buffer quality, invasive plant growth, basal area of trees, and presence or absence of stressors were noted. Within the stream portion of the sampling area, stream incision was measured, presence or absence of stressors were noted, and a stream habitat assessment was conducted. SWR Assessments for this project took place 23-24 June 2014.

Intensive Level: Macroinvertebrate Index of Biotic Integrity

Collection, processing, and analysis of benthic macroinvertebrates followed Pennsylvania Department of Environmental Protection (PA-DEP) protocol for wadeable, freestone riffle-run streams (2009). PA-DEP protocol defines two general sampling seasons: November through May, and June through September. However, the June through September window is not recommended for macroinvertebrate sampling because many macroinvertebrates are in the egg stage of development during the summer and would not be collected, reducing the strength of IBI metrics. Therefore, sampling took place in the macroinvertebrate preferred sampling window during May 2014. There are also two stream size categories according to the PA-DEP protocol: first to third order streams (Strahler 1952) that drain less than 65 km^2 (25 mi^2) of land or large fifth and higher order streams that drain more than 130 km^2 (50 mi^2) of land. As the study watershed is 7.3 km^2, all sites in this study fell into the former category.

Collection was conducted using a D-frame net with 500 µm mesh in a 100-m stream reach. Working upstream, six kicks were collected from a variety of riffle habitat (e.g., slow/shallow and fast/deep) by disturbing one square meter of substrate for one minute to a depth of 10 cm immediately upstream of the net for each kick. Riffle samples were composited for each site and preserved in 95% ethanol in the field.

Samples were transported to the laboratory for sorting and identification. First, a subsample of 200 ± 40 organisms was separated out for each composited sample, with the preferred range of 190-210 total identifiable organisms. This was done by placing the collected sample in a 35.6-cm x 20.3-cm x 8.9-cm (14” x 8” x 3.5”) pan with 28 – 5.1-cm x 5.1-cm (2” x 2”) grid cells. Four random cells were selected and moved to a second pan, where all organisms were picked. If less than 160 identifiable organisms were picked, additional cells were randomly

selected, one at a time, until the final subsample size fell within the preferred range. If more than 240 identifiable organisms were picked in first four cells, this subsample was placed in second pan and subsampled to 200 ± 40 organisms. The total number of cells picked was recorded.

All organisms were then identified and counted using standard taxonomic keys (e.g., Merritt et al. 2008, Stewart et al. 2002, Peckarsky 1990). Identification was taken to the genus level for most individuals except:

- Phylum level: roundworms (Nematoda) and proboscis worms (Nemertea), moss animalcules (Bryozoa)
- Class level: flatworms (Turbellaria), leeches (Hirudenia), and segmented/aquatic earthworms/tubificids (Oligochaeta)
- Water mites classified as Hydracarina, an artificial taxonomic grouping of several water mite superfamilies
- Family level: chironimids, snails, clams, and mussels

A habitat assessment was also conducted at each site at the time of sampling in accordance with EPA's Rapid Bioassessment Protocol (Barbour et al. 1999). The "physical characterization and water quality" data sheet as well as the "habitat assessment field data sheet for high gradient streams" was completed in the field. These field worksheets are publicly online from the EPA at http://water.epa.gov/scitech/monitoring/rsl/bioassessment/. After parameters were evaluated in the field, numerical ratings were summed to provide a habitat ranking (Barbour et al. 1999).

The PA-DEP IBI incorporates a suite of metrics that measure biological attributes that respond to stressors (Table 2). This response is based in part on Pollution Tolerance Values (PTVs) assigned by PA-DEP, ranging from zero to 10, that mostly reflect response to organic

enrichment and sedimentation. There are six core metrics in the PA-DEP IBI: Total Taxa Richness, Ephemeroptera-Plecoptera-Trichoptera (EPT) Taxa Richness, Beck's Biotic Index (Version 3), Shannon Diversity Index, Hilsenhoff Biotic Index (HBI), and Percent of Sensitive Individuals. As stream stressors increase, component metric scores decrease except for HBI, whose scores increase with increasing stress. The overall IBI score decreases with increasing stressors.

Table 2. The six component metrics to the PA-DEP Index of Biotic Integrity and descriptions of what each metric measures. EPT refers to the entomological classes Ephemeroptera, Plecoptera, and Trichoptera. PTV refers to Pollution Tolerance Values assigned by PA-DEP.

Metric	**What it Measures**
Total Taxa Richness	Number of taxa
EPT Taxa Richness	Number of taxa in E, P, or T. Only includes those taxa with PTVs ≤ 4
Beck's Index (Version 3)	Weighted count of taxa with PTV ≤ 2
Shannon Diversity	Richness and evenness
Hilsenhoff Biotic Index	Average number of individuals weighted by PTVs
Percent Sensitive Individuals	Percent of individuals with PTV ≤ 3

Additional Empirical Measurements

In addition to the ecological assessments, chemical and physical parameters were also measured at each sampling location. The pH, conductivity, specific conductivity, dissolved oxygen, and temperature were measured in the field using handheld probes. Measurements were planned to be taken on the day of macroinvertebrate sampling for each site, but due to a damaged pH bulb they were postponed and all completed on 11 June 2014.

Grab samples for nitrate and ammonium were collected once in the summer of 2014, once in the fall of 2014, and once in the spring of 2015 during baseflow conditions. Samples were collected in 60 mL containers, stored on ice in the dark, and transported to the USDA-ARS

facility in State College, PA. Here, samples were vacuum filtered and transferred to the USDA-ARS chemical analysis laboratory for nitrate and ammonia analyses. Field and lab blanks were also included in the analyses.

The nitrate and ammonia analyses were conducted using the QuickChem® Method 10-107-04-1-R (Harbridge and Tucker 2008) and QuickChem® Method 10-107-06-2-L (Harbridge 2007) protocols, respectively. For nitrate measurements, the samples were passed through a copperized cadmium column to first reduce it to nitrite. A reaction with sulfanilamide followed by coupling with N-(1-naphthyl) ethylenediamine diydrochloride to produce a magenta soluble dye that was read using a colorimetric assay. For ammonia measurements, samples were heated with salicylate and hypochlorite in an alkaline phosphate buffer. The reaction produced an emerald green product that was intensified with sodium nitroprusside. A colorimetric assay was then used to measure ammonia concentration.

Soil and Water Assessment Tool

The SWAT modeling in this study was done by Amy Collick, Ph.D. and Tamie Veith, Ph.D. of the USDA-ARS Pasture Systems and Watershed Management Research Unit in State College, PA. The following summary of SWAT set-up, data sources, and application was prepared by Amy Collick, Ph.D. (Personal Communication, July 7, 2015):

Method

To compare the ecological indexing values with modeled nutrient output, TopoSWAT was initialized in WE38 with small subbasins at each of the stream sites where the indices were determined.

Input Data

Spatial Data

A 10-m digital elevation model (DEM) resampled from a finer resolution LiDAR (< 1m) ((Canaan Valley Institute, Morgantown, West Virginia, 2007)) projected in NAD1983, UTM 18N was the basis of all the topographic information necessary for TopoSWAT. Following the method established in Easton et al. (2008), TopoSWAT created the soils layer from the intersection of the topographic wetness index (TI) layer with a map of soils including the associated soil properties from the Food and Agriculture Organization (FAO-UNESCO) Digital Soil Map of the World (Land and Water Development Division, FAO, Rome, 2007; http://www.fao.org/geonetwork/srv/en/main.home#). A soils lookup table was also generated. The land use spatial layer was composed of more than 300 farm fields digitized in ArcGIS (ESRI 2012. ArcGIS Desktop: Release 10.1 Redlands, CA: Environmental Systems Research Institute) using annually updated geographical positioning system (GPS) surveys of the field boundaries then aggregated by common crop rotations and field management (Trimble Navigation, Limited 2015).

Field Management

Planting dates and crops, harvest dates and yields, tillage practices, and fertilizer or manure applications were recorded in annual field operation histories. These histories were collected each winter via in-person interviews between the farmers and a single USDA-ARS field technician over a six-year

period for most fields in WE38 and over a 12-year period for fields in the experimental subwatersheds near sampling Site 8 and upstream of Site 7 (Veith et al. 2015). Using these data, cropping rotations and associated management schedules were developed for each field in WE38, formatted and coded for SWAT, and spatially represented in SWAT as distinct fields, each comprised of HRUs of various TI classes.

Weather, Hydrology, and Water Quality Data

Daily averages were calculated from five-minute interval datasets of precipitation, minimum and maximum temperatures, relative humidity, and wind speed collected from 1986 to 2010 at a USDA-ARS weather station within the watershed. Continuous discharge and water quality variables (phosphorus and nitrogen) from the 1980s to 2010 were used to corroborate the model outputs of stream flow and nutrient transport.

Model Set Up

The 14 index sampling points in WE38 were used as a reference to mark outlet points 50 m upstream and 50 m downstream of each sampling point to establish a subbasin for each sampling transect. Thus, 29 subbasins were established (Figure 6).

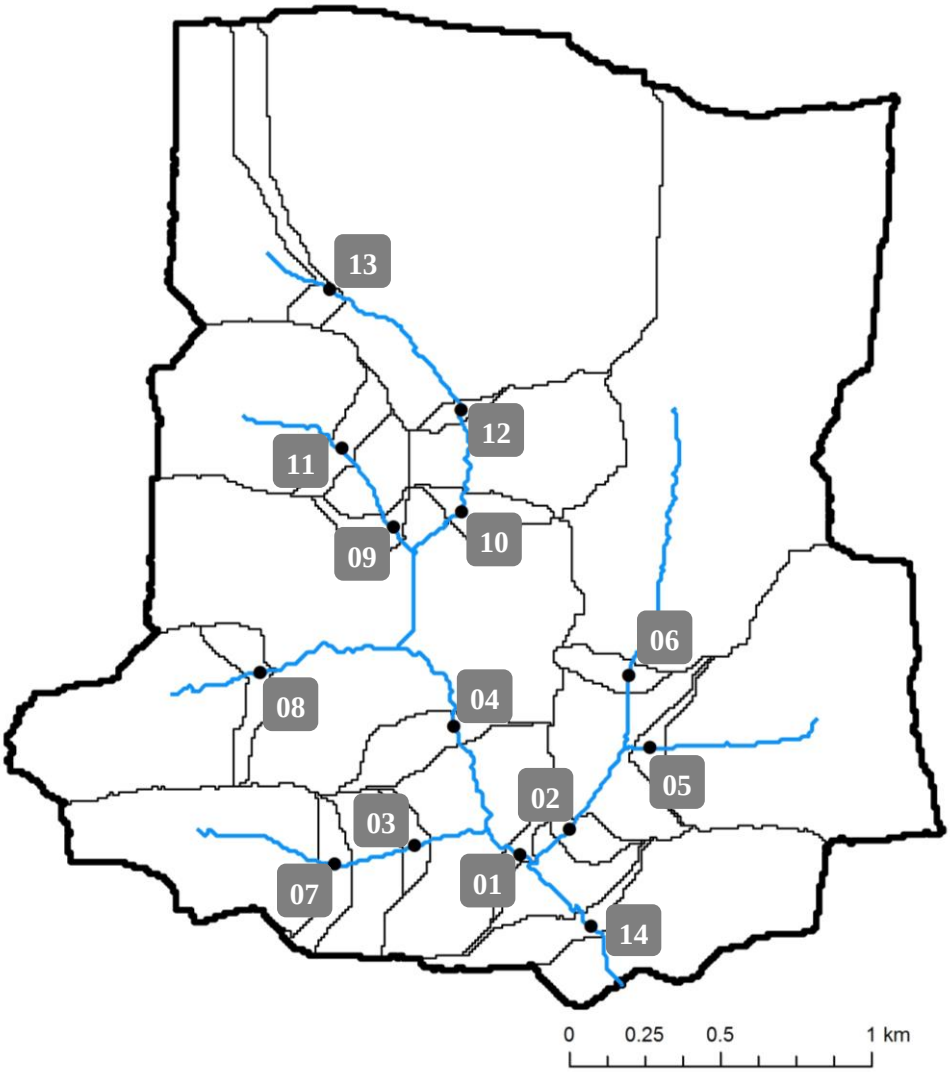


Figure 6. Sampling sites and transect subbasins delineated in TopoSWAT and shown as black lines; streams are shown as blue lines.

HRUs were generated by intersecting spatial layers of land use, TI-classed soils, and hydrologic subbasins. Since the land use was divided into farm fields, unique HRUs were generated within each field. Therefore, the detailed field management practices could be specified for each HRU within each of the separate fields with different crop rotations and practice schedules. The

calibration parameters established in Collick et al. (2014) were used to parameterize the current model. The Nash-Sutcliffe efficiency and visual inspection of the hydrograph and water balance were used to confirm that the model performed as well as in Collick et al. (2014).

Model Application

The model was run for 24 years from 1987 to 2010 on a daily time step, and monthly output from 1999 to 2010 was used for analysis. Output variables pertaining to water yield, flow, sediment yield, phosphorus (soluble or dissolved, sediment-bound, organic, and total) loss, and nitrogen (organic, nitrite, and ammonium) loss were extracted from the Reach, Subbasin, and HRU output files.

The first research question pertained to identifying the available scales of SWAT outputs. This was done through examination of output files and extensive meetings and discussions with Drs. Collick and Veith, as well as a consideration of the tools available in GIS applications to create additional analysis scales from the output post-modeling. All desired raw model output files were then provided by Drs. Collick and Veith.

Variable Reduction

The second research question of this project asked which of the identified scales of SWAT output are redundant for the objectives of this research, with the aim of eliminating redundant variables and identifying unique spatial and temporal scales of SWAT output that could be used for further comparisons with ecological data. SWAT variables were compared to

one another using Spearman's rho Rank-Order Correlation statistic to generate a correlation matrix using Minitab 17 (Minitab Statistical Software, 2010. State College, PA: Minitab, Inc.). Variables that were highly correlated ($p<0.05$, $r>0.7$) were grouped and a single representative variable was chosen from each group to be used in the proceeding analyses. This was achieved by examining the r and p values in each correlation matrix to identify the single variable having the strongest and highest number of correlations to the other variables in each grouping. This allowed the greatest variable reduction possible and selected for representative variables that conveyed nutrient runoff information that was comparable to the other variables in their groupings. The final selection of SWAT variables was then analyzed again to verify that they were not significantly correlated ($p>0.05$, $r<0.07$), which was a requirement for regression.

To decrease the number of variables from the landscape, rapid, and intensive ecological assessments, correlation matrices were also calculated among the metrics within each of these groups. Each group was analyzed separately to select a small list of representative variables for each of the ecological assessment groups ($p<0.05$, $r>0.7$). The IBI component metric comparisons were conducted using Pearson's Correlation (Minitab 17) because all component metrics and the overall IBI score were either normally distributed or transformed to achieve normality. The other comparisons did not achieve normality and, therefore, Spearman's rho Correlation was used.

Multiple Regression

An iterative Best Subsets Regression (Minitab 17) was performed on the final variables to describe the relationship between SWAT nutrient output and ecological condition in WE38. The first iteration included the final (i.e., representative) SWAT output variables as independent

predictor variables and the IBI score as the dependent response variable. The strongest model or models were determined based on relative changes in R^2, R^2 adjusted, R^2 predicted, Mallows C_p, and S with one-at-a-time variable additions. The SWAT variables in the strongest model or models (i.e., most well-fit model across all statistical measures) were kept for the next step, while less related SWAT variables were eliminated.

The second iteration included these strongest SWAT variables plus the representative percent land use variables, after parameter reduction as described in the previous section, in the Best Subsets Regression of IBI. This was to assess whether SWAT, percent land use, or a combination of both were most informative in explaining IBI scores. Again, the best model was determined by considering overall best fit based on relative changes in the five metrics listed above as variables were added into the regression. Less explanatory variables were then eliminated for the next regression.

The third iteration included these strongest SWAT and percent land use variables, but the representative Stream-Wetland-Riparian Index variables were also added to the Best Subsets Regression of IBI. This was to assess whether SWAT, percent land use, SWR, or a combination were most informative in explaining IBI scores. Again, the best model was determined by considering the five metrics listed above.

A comparison of the best models from the three iterations of Best Subsets regression was then made to determine the strongest overall model. Conclusions were then drawn based on the variables found to be most explanatory and predictive of IBI score.

Additional Comparisons

Analyses of the grab sample nutrient concentrations were also evaluated to describe the relationship between grab sample nutrients and SWAT-modeled nutrients. Statistical significance was determined using Spearman's rho ($p<0.05$) (Minitab 17). Additionally, a Best Subsets Regression was performed with the variables from the best regression model above. This was used to assess the explanatory value of grab samples for IBI Score as compared to SWAT, LULC, and SWR measurements. Again, R^2, R^2 adjusted, R^2 predicted, Mallows C_p, and S were used to compare models.

RESULTS

SWAT Output Scales

The SWAT model produced three main output files: Subbasin, Reach, and HRU (Hydrologic Response Unit). In total, six spatial scales were described from SWAT model outputs: two from SWAT output directly and four from post-processing of SWAT data.

The "Subbasin" output file provides data for each subbasin polygon that is designated in model set-up. In this project, it provided nutrient runoff data for each of the 29 subbasins in WE38. Because the modeling team designated outlets at both the upstream and downstream endpoints of each 100-m sampling reach transect, there is a subbasin polygon associated with each 100-m sampling site. This means that each site-level subbasin (i.e., the narrower polygons in Figure 7a-b) includes only the land area that drains directly to that 100-m reach; anything that drains upstream of the 100-m sample reach is included in the wider polygons of Figure 7a-b. By looking at only these 14 narrower subbasins (one for each sample site), this output forms the first spatial scale, "Subbasin-Site Only" (Figure 7a).

The second spatial scale also comes from the "Subbasin" output file, but includes the nutrient runoff generated from any land within the entire upstream contributing area. This was calculated by combining nutrient runoff from all polygons within and upstream of the sample reach. Figure 7b shows an example of the subbasins that would be included in the "Subbasin-All Upstream" scale for Site 10. In comparison, Figure 7a depicts the single subbasin that would be included in the "Subbasin-Site Only" scale for the same site.

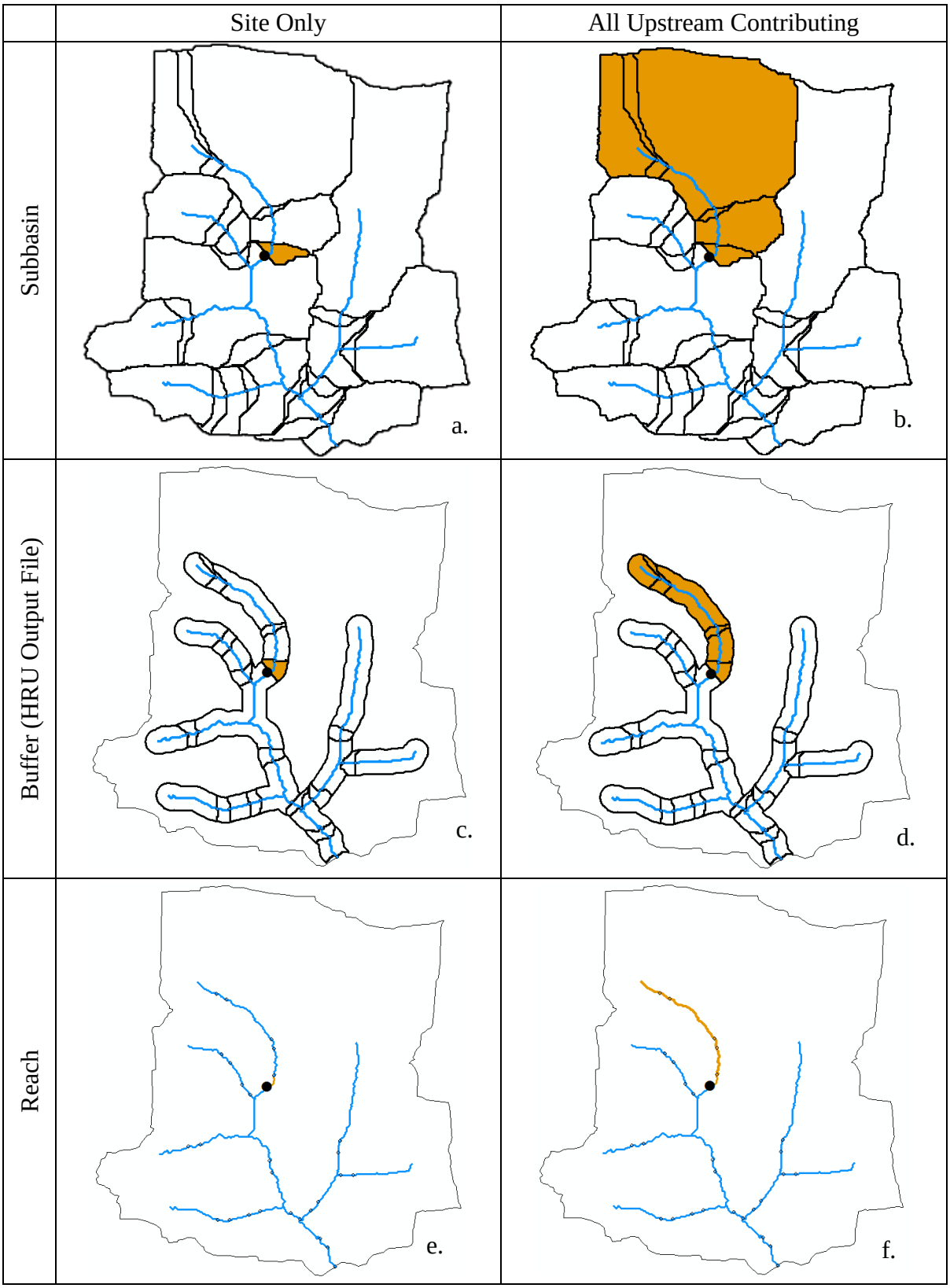

Site Only
All Upstream Contributing
Subbasin
Buffer (HRU Output File)
Reach
a.
b.
c.
d.
e.
f.

Figure 7. (Previous page) The six spatial scales of SWAT output considered in this study. Columns describe the longitudinal aspect of the contributing area: Site Only and All Upstream. Rows describe the lateral aspect of the contributing area: Subbasin, Buffer, and Reach. Each column-row combination is depicted with an example of the data included at each scale. Blue lines represent streams in the WE38 watershed. Black polygons represent the subbasins delineated in the SWAT model. Orange shading represents the SWAT output included in that spatial scale, using Site 10 as an example.

The "HRU" output file provides data for each HRU polygon that is identified from the Topographic Index (TI) and Land Cover layers in ArcSWAT. In this study, over 3,000 HRUs were identified in the WE38 watershed, so the "HRU" output file provided nutrient runoff data for all 3,000+ HRUs within WE38. Unlike the "Subbasin" output file, the raw data in the HRU output file are not directly relevant to instream ecological studies. However, the file can be processed in a GIS to produce two pertinent scales: the "Buffer-Site Only" and "Buffer-All Upstream."

As stated previously, the SWAT model does not route runoff across HRUs to the stream, but instead assumes that all runoff generated in each HRU is transported without processing to the subbasin outlet. In the landscape, runoff generated farther from the stream will not have this direct connection to the stream outlet, but rather the volume and nutrient loads will likely decrease or may never reach the outlet. In TopoSWAT, the HRUs in close proximity to the stream are likely to be classified in higher TI classes, thus contributing the most significant portion of runoff to the stream reaches. The "Subbasin" scale, however, does not differentiate HRUs and treats them collectively; therefore, the "Buffer" scales were created from the "HRU" output file to include only those HRUs within a 100-m buffer on either side of the stream reach. As with the "Subbasin" output, there are two contributing scales: "Buffer-Site Only" includes only those HRUs within the 100-m buffer that contribute directly to the sample reach (Figure

7c); “Buffer-All Upstream” includes only those HRUs within the 100-m buffer that contribute anywhere in or upstream of the sample reach (Figure 7d).

The “Reach” output file provides data for each stream reach. In this study, it provided data for each of the 29 stream reaches (corresponding to each of the 29 subbasins) in WE38. For “Reach” data, SWAT models instream nutrient concentrations by adding runoff coming from the “Subbasin-Site Only” scale with runoff coming from all upstream subbasins, except that this upstream runoff is further modeled to include instream nutrient processing. In this way it is better able to account for the proportion of nutrients that enter the stream network in the upper-most reaches of headwater tributaries and are taken up or transformed before they arrive at stream reaches farther downstream. As the “Buffer” scales attempt to address lateral distance decay, the “Reach” scales attempts to address longitudinal distance decay. As with the other scales, there are two contributing scales for “Reach:” the “Reach-All Upstream” is the scale provided in the Reach output (Figure 7f). The “Reach-Site Only” was generated in post-processing by subtracting the “Reach-All Upstream” runoff from the reach directly above a sampling reach from the “Reach-All Upstream” runoff within that sampling reach (Figure 7e).

In addition to the six spatial scales described, temporal scales were also considered. The SWAT model was run on a daily timestep and data were generated for a 12-year period (1999-2010). This provided an immense amount of runoff data to compare with ecological metrics. Ultimately five temporal scales were selected for this study: average annual and average seasonal (for each of winter, spring, summer, and fall). Average annual was a mean nutrient concentration of the entire year and across all years of the study. Seasonal averages provided a mean nutrient concentration for a three-month portion of the year: winter (January – March), spring (April – June), summer (July-September), and fall (October – December). The annual

scale was included to provide a broad overview of the nutrient conditions of the stream reach. The seasonal scales were included because some of the details of nutrient runoff patterns may be lost in averaging across the whole year. Because Pionke et al. (2000) found for this watershed that the majority of nitrate runoff occurred in winter and spring and that the majority of phosphorus runoff occurred with the most severe storms, these seasonal differences could prove important. In addition to seasonal runoff trends, macroinvertebrate ecology varies and certain life stages may be more or less susceptible to nutrient stress.

In all, nutrient runoff data for each of the five temporal scales were calculated for each of the six spatial scales, yielding 30 total spatio-temporal scales of SWAT output. Three nutrients were included for each of these scales: Total Phosphorus (mg/L), Total Nitrate (mg/L), and Total Organic Nitrogen (mg/L). This resulted in 90 total SWAT variables to consider (Table 3).

Table 3. Spatial and temporal scales of SWAT model output for nutrient runoff.

Major Spatial Scale		**Minor Spatial Scale**		**Temporal Scale**		**Nutrient**
Subbasin	**x**	All Upstream Contributing	**x**	Annual	**x**	Total Phosphorus (mg/L)
100-m Buffer		Site Only		Winter		Total Nitrate (mg/L)
Reach				Spring		Total Organic Nitrogen (mg/L)
				Summer		
				Fall		

SWAT Variable Reduction by Spatio-Temporal Scale

The aim of the SWAT variable reduction was to eliminate redundant spatio-temporal scales of modeled nutrient runoff and to develop a condensed list of unique variables to use in a model for biological integrity using IBI score. In the end, the pool of 90 variables was

condensed to only five (Table 4). There was a high level of redundancy and these final five variables represented 51 of the original 90 spatio-temporal SWAT variables (Table 5).

Table 4. The final five representative SWAT variables after Round Two elimination.

SUSuN3	Subbasin-All Upstream Average Summer Nitrate
SUFaN3	Subbasin-All Upstream Average Fall Nitrate
SSWP	Subbasin-Site Only Average Winter Phosphorus
SSSuN3	Subbasin-Site Only Average Summer Nitrate
RUSpP	Reach-All Upstream Average Spring Phosphorus

Table 5. The final five representative SWAT variables (in bold) and SWAT variables with which they are correlated (Spearman's rho, p<0.05). The first letter codes R for Reach, S for Subbasin, or B for Buffer. The second letter codes U for All Upstream or S for Site Only. The third letter group conveys temporal scale: W, Sp, Su, or F for Annual, Winter, Sp, Su, or Fall, respectively. The last letters code for nutrient: P for Phosphorus, N3 for Nitrate, or ON for Organic Nitrogen.

RUSpP	**SSWP**	**SSSuN3**	**SSSpN3**	**SUFaN3**
BUAnN3	SSAnP	BSAnN3	BSAnN3	BUAnP
RUAnP	SSAnON	RUSuN3	BSAnON	BUAnN3
RUAnN3	SSWON	SSAnN3	RUAnN3	BUAnON
RUAnON	SSSpP	SSWN3	RUAnON	RUAnP
RUWP	SSSuP	SSFaN3	RUWN3	RUSpP
RUWN3	SSFaP		RUSpN3	RUSuP
RUWON	SSFaN3		RUSpON	RUFaP
RUSpN3	SSFaON		RUSuN3	RUFaN3
RUSpON	BSAnP		RUSuON	SUAnP
RUSuP	BSAnON		RUFaN3	SUAnN3
RUSuN3			RUFaON	SUAnON
RUSuON			SSAnN3	SUWP
			SSAnON	SUWN3
			SSWN3	SUWON
			SSWON	SUSpP
			SSSpON	SUSpN3
			SSSuON	SUSpON
			SSFaON	SUSuP
				SUSuN3
				SUSuON
				SUFaP
				SUFaON

The remaining 39 of 90 variables that were not represented were eliminated from further consideration early during the research process. First, all "Reach-Site Only" variables (Figure 7e) were excluded because the method of subtracting the upstream reach runoff from the sample reach runoff did not function as expected. Some values for nutrients and water yield were negative, perhaps due to the mathematical procedures modeling instream processing being more complicated than originally understood. Further investigation into the model is necessary to better understand this issue and develop a way to isolate the "Reach-Site Only" data.

Additionally, the "Buffer" seasonal scales for both "Site Only" and "All Upstream" were also excluded. To convert 3,000+ HRUs from monthly averages to seasonal averages and then isolate only those HRUs within the 100-m buffer proved to be an immense amount of processing. Thus, it was determined to proceed using only the average annual temporal scale for the two "Buffer" scales and then revisit seasonal averages if annual results suggested they would provide additional information. This did not prove true, however, as is described in the subsequent section, so the seasonal scales were eliminated from further analysis.

Accounting for the "Reach-Site Only" and "Buffer" seasonal variables, this left 51 of the original 90 SWAT variables. Because all variables are derived from the same data set, there was a possibility of redundancy among them. The first round of elimination using Spearman's rho correlation was among variables within each of the six spatial scales individually. Correlation matrices were generated (one for each spatial scale) comparing differences among temporal scales and nutrients. However, within each spatial scale, nutrient runoff was frequently correlated across both temporal scales and nutrients ($p<0.05$, $r>0.06$). That is, the annual, winter, spring, summer, and fall variables all provided the same patterns of nutrient runoff in WE38 for a particular spatial scale. Total phosphorus, nitrate, and organic nitrogen were generally correlated

with one another as well. Next, from each spatial scale, all the variables were categorized into unique (i.e., non-correlated) groupings. Thus, there were two unique groupings of variables from "Subbasin-All Upstream," three groupings for "Subbasin-Site Only," one grouping from "Reach-All Upstream," one grouping from "Buffer-Site Only," and one grouping from "Buffer-All Upstream." A single variable was selected to represent each of these eight variable groups, and these eight variables are listed in Table 6.

Table 6. The eight representative variables remaining after Round 1 correlation elimination of SWAT variables.

SUSuN3	Subbasin-All Upstream Average Summer Nitrate
SUFaN3	Subbasin-All Upstream Average Fall Nitrate
SSWP	Subbasin-Site Only Average Winter Phosphorus
SSSuN3	Subbasin-Site Only Average Summer Nitrate
SSSpN3	Subbasin-Site Only Average Spring Nitrate
RUSpP	Reach-All Upstream Average Spring Phosphorus
BSAnON	Buffer-Site Only Average Annual Organic Nitrogen
BUAnN3	Buffer-All Upstream Average Annual Nitrate

A second round of correlations using Spearman's rho assessed redundancy of these eight representative SWAT variables. Because the first round of correlations only looked at redundancy of variables within each spatial scale (e.g., comparing "Reach-All Upstream" variables to each other only), the second round compared variables across spatial scales (e.g., comparing "Reach-All Upstream" with "Buffer-Site Only" variables). "Buffer-Site Only Annual Organic Nitrogen" was significantly correlated with "Subbasin-Site Only Spring Nitrate" ($r=0.72$, $p=0.002$). "Buffer-All Upstream Annual Nitrate" was significantly correlated with "Subbasin-All Upstream Fall Nitrate" ($r=0.89$, $p<0.001$) and "Reach-All Upstream Winter Phosphorus" ($r=0.75$, $p=0.001$). Both Buffer variables were thus eliminated during round 2

elimination. "Subbasin-All Upstream Summer Nitrate" was correlated with "Reach-All Upstream Spring Phosphorus" (r=0.77, p=0.001) and "Subbasin-Site Only Spring Phosphorus" (r=0.62, p=0.014) and was, therefore, eliminated as well. After excluding these variables, five unique groups of SWAT variables remained. These variables are listed in Table 4. These five SWAT variables were the final unique variables to be tested against the ecological metrics. The other SWAT variables that these final five represent (i.e., with which they correlate) are listed in Table 5.

Ecological Assessments

Landscape Assessment

After identifying the spatial scales represented by the SWAT output, landscape scales were chosen to relate as closely as possible to the six SWAT scales. Land use in the "Subbasin-Site Only", "Subbasin-All Upstream," "Buffer-Site Only," and "Buffer-All Upstream" scales corresponded to the same SWAT land use scales. The land use included in the SWAT "Reach" scales matches with the land use in the Subbasin scales, so separate "Reach" land use variables were not included. Two additional spatial scales were also included: a 100-m landscape circle and a 1-km landscape circle. Landscape circles are easily generated within a GIS by buffering around a sample point to a designated radius. In larger watersheds, 1-km is often typical. Because this is a smaller watershed, a 100-m radius was also included. A summary of the six landscape-level scales is shown in Table 7.

Table 7. Spatial scales of analysis for land cover classes.

Spatial Scale		**Land Cover Variable**
Subbasin-All Upstream	**x**	Percent Forest
Subbasin-Site Only		Percent Agriculture
100-m Buffer-All Upstream		Percent Urban
100-m Buffer-Site only		
1-km Circle		
100-m Circle		

The percentage of forest, agriculture, and urban land cover was calculated for each of the 14 sample sites in this study. It was found that, within each of the six spatial scales, forest and agriculture were always highly correlated (Table 8). Thus, the Percent Agriculture variables were removed from further analysis due to redundancy. Additionally, because the WE38 watershed has very low urban development, the Percent Urban variables were also removed from further analysis due to irrelevancy. The patterns of Percent Forest for each of the six scales are show in Figure 8.

Table 8. Regression of Percent Agriculture and Percent Forest within each spatial scale. All comparisons were found to be significant.

REGRESSION: Agriculture versus Forest						
Statistic	*100-m Circle*	*1-km Circle*	*Subbasin-Site*	*Subbasin-Up*	*Buffer-Site*	*Buffer-Up*
p	0.000	0.000	0.000	0.000	0.000	0.000
R^2	93.6%	99.3%	95.3%	97.3%	94.4%	86.0%
R^2-adj	93.1%	99.2%	94.9%	97.0%	93.9%	84.9%

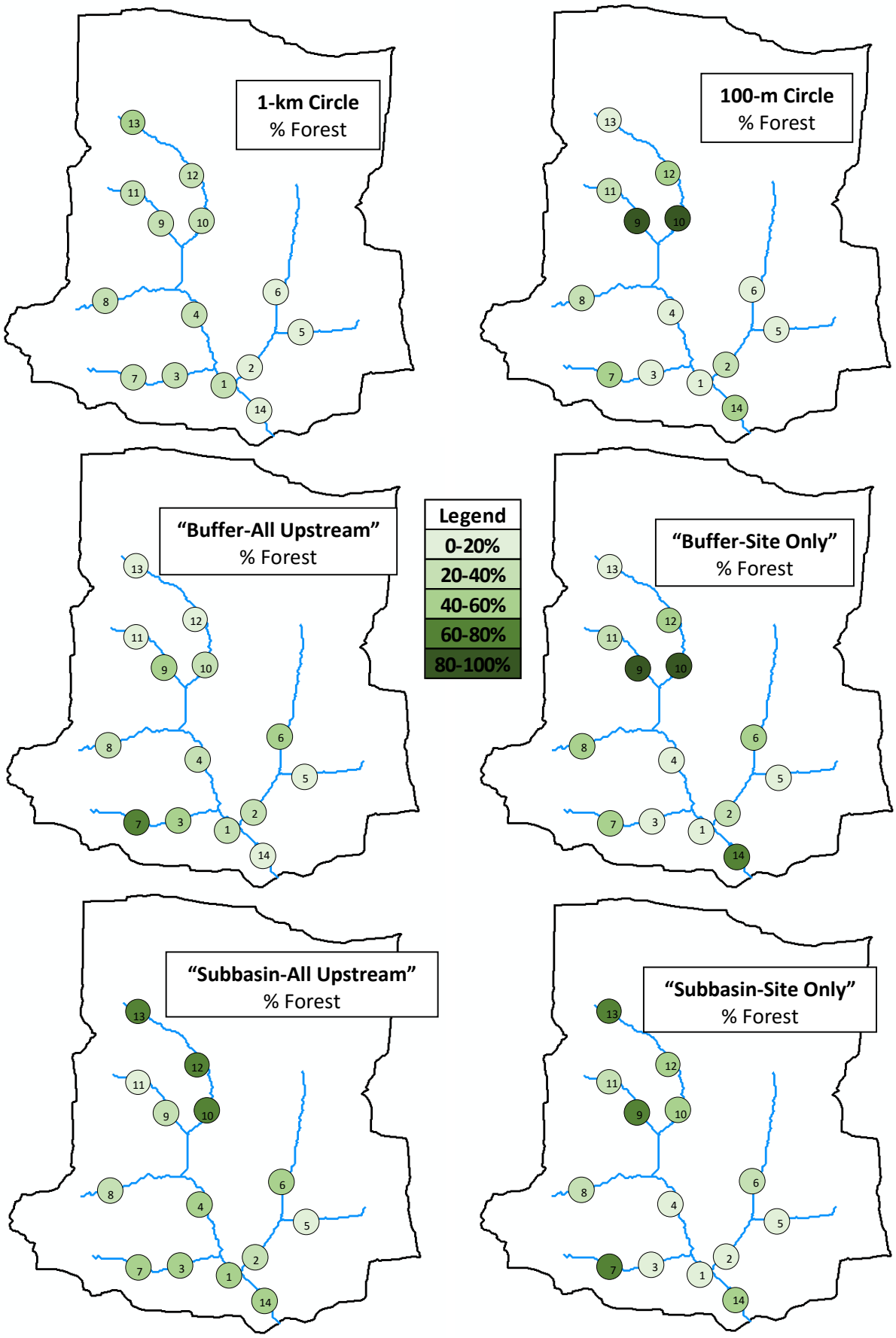

1-km Circle
% Forest
100-m Circle
% Forest
"Buffer-All Upstream"
% Forest
Legend
0-20%
20-40%
40-60%
60-80%
80-100%
"Buffer-Site Only"
% Forest
"Subbasin-All Upstream"
% Forest
"Subbasin-Site Only"
% Forest

Figure 8. (Previous page) Percent Forest for each sampling site at six landscape scales. Circles are labeled with the Site ID number and shaded to represent Percent Forest at each scale.

The correlation matrix among the remaining six variables (Percent Forest for the six spatial scales) revealed that the 100-m Circle, "Buffer-Site Only," and "Subbasin-Site Only" scales were correlated (Spearman's rho, $r>0.57$, $p<0.05$) (Table 9). Percent Forest in "Subbasin-Site Only" was selected as the representative variable of these three spatial scales; 100-m Circle and "Buffer-Site Only" were eliminated from further analyses. Additionally, examination of the 1-km circle Percent Forest values showed that this spatial scale is too broad for this small watershed. The 1-km radius landscape circles created for each sample point in ArcGIS overlapped substantially with one another, and the result was a dilution of relevant landscape-level details (Figure 8). Thus, the 1-km landscape circle variables were eliminated from further analyses. The remaining unique landscape-level variables were the Percent Forest in the "Subbasin-Site Only," "Subbasin-All Upstream," and "Buffer-All Upstream" scales.

All of the three remaining land use variables described a wide range of Percent Forest values across the WE38 watershed. According to the "Subbasin-Site Only Percent Forest" scale, sampling points ranged from 0-65% forest, with a mean of 33%. "Subbasin-All Upstream Percent Forest" ranged from 4%-75%, with a mean of 43% forest. "Buffer-All Upstream Percent Forest" ranged from 0%-62%, with a mean of 30% forest.

Table 9. Correlation matrix of Percent Forest. The 1-km landscape circle scale was removed due to its irrelevancy. The 100-m landscape circle, "Buffer-Site Only," and "Subbasin-Site Only" scales were found to be significantly correlated.

		100-m Circle Percent Forest	**Subbasin-Site Only Percent Forest**	**Subbasin-All Upstream Percent Forest**	**Buffer-Site Only Percent Forest**
Subbasin-Site Only Percent Forest	r	0.575			
	p	0.025			
Subbasin-All Upstream Percent Forest	r	-0.004	0.329		
	p	0.99	0.232		
Buffer-Site Only Percent Forest	r	0.822	0.607	0.079	
	p	0	0.016	0.781	
Buffer-All Upstream Percent Forest	r	0.222	0.2	0.325	0.396
	p	0.427	0.475	0.237	0.143

Stream-Wetland-Riparian Index Score

The Stream-Wetland-Riparian Score component metrics and overall score were calculated for the 14 sample sites. There were eight component metrics calculated: Buffer score, basal area score, floodplain-wetland score, floodplain-wetland condition score, stream habitat assessment score, incision, stressors, and invasive plants score. These were combined to form the SWR Score for each site, as shown in Figure 9. The SWR score ranges from 0 to 1, with a lower score indicating a more impacted condition. Scores in WE38 ranged from 0.36 to 0.79, with an average of 0.59.

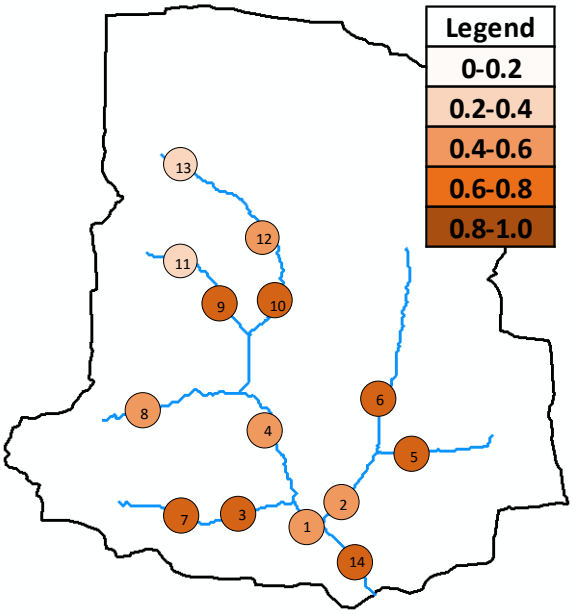


Figure 9. Stream-Wetland-Riparian (SWR) Index Score for each sampling site. Circles are labeled with Site ID and shaded to represent SWR Score.

The invasive plants score was found to be uncorrelated with all other SWR variables, highly non-normal, and lacking information. This is likely due to the livestock access to streams and riparian mowing that occurred in sampling sites and precluded invasive growth as a predictor of stream, wetland, and riparian condition. Thus, invasive score was removed as an individual variable from further analyses, though it did remain a part of the overall SWR score.

The correlation matrix identified two unique groupings of SWR variables: those that correlated with overall SWR Score (Table 10) and those that correlated with Floodplain-Wetland Score (Table 11). Thus, overall SWR Score and Floodplain-Wetland Score were chosen to represent the other SWR variables in further analyses.

Table 10. Correlation matrix of overall SWR Score with component metrics (Invasive Score excluded). Significant correlations (p<0.05) were seen with Buffer, Floodplain-Wetland (FPWL) Condition, Stream Habitat Assessment (SHA), Incision, and Stressor Scores.

		Buffer	Basal	FPWL	FWLCond	SHA	Incision	Stressors
SWR	**r**	0.536	0.409	0.503	0.635	0.881	0.651	0.817
	p	0.048	0.146	0.067	0.015	0	0.012	0

Table 11. Correlation matrix of Floodplain-Wetland (FPWL) Score with component metrics (Invasive Score excluded). Significant correlations (p<0.05) were seen with Basal Area Score and Floodplain-Wetland Condition Score.

		Buffer	Basal	FWLCond	SHA	Incision	Stressors	SWR
FPWL	**r**	0.468	0.835	0.829	0.369	0.035	0.183	0.503
	p	0.091	0	0	0.195	0.904	0.531	0.067

Macroinvertebrate Index of Biotic Integrity

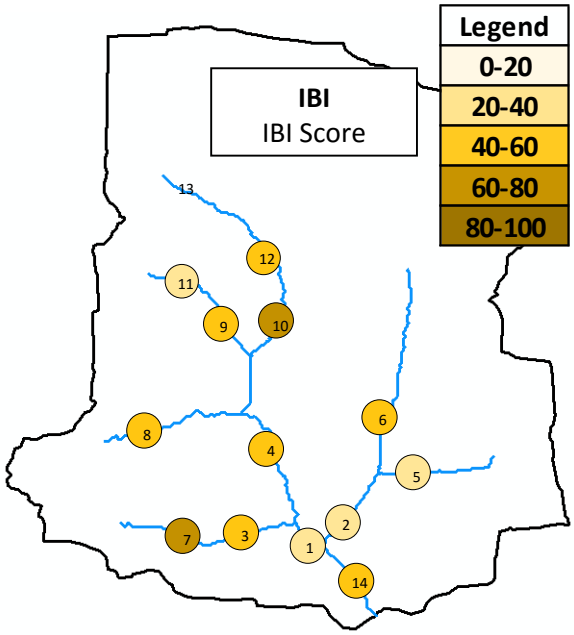


Figure 10. Index of Biotic Integrity (IBI) Scores. Circles are labeled with Site ID number and shaded to represent IBI Score.

The Index of Biotic Integrity (IBI) component metrics and overall score were calculated for 13 of the 14 sample sites (Figure 10). Note that Site 13 was excluded from analyses because it did not meet the perennial stream IBI requirement set forth by the Pennsylvania Department of Environmental Protection (PA-DEP). Macroinvertebrates at Site 13 were collected, sorted, and identified during the spring sampling, but during collection of summer grab samples the stream was found to have dried. All other streams remained flowing through the year and are thought to fulfill the perennial stream prerequisite. Samples were collected, subsampled, and identified according to PA-DEP guidelines, resulting in >2,600 macroinvertebrates that were identified for use in the IBI.

The range of IBI scores was broad for such a small, agriculturally-dominated watershed. The lowest IBI score was 20.5, indicating highly impacted conditions. The highest IBI score was 79.3, reflecting a relatively high-integrity site. The mean IBI score for the WE38 watershed was 48.9, indicating that many of the sampled stream reaches experienced degraded biological integrity.

There were six component metrics to the IBI: Total taxa richness; Ephemeroptera, Plecoptera, and Trichoptera Richness; Beck's Index; Hilsenhoff Biotic Index; Shannon Diversity Index; and Percent Sensitive Individuals. HBI and Percent Sensitive Individuals were natural log transformed to achieve normality. A correlation matrix of all IBI variables using Pearson correlation showed that all six component metrics were significantly correlated with overall IBI Score ($r>0.64$, $p<0.02$ for all variables) (Table 12). Therefore, the overall IBI score was the only variable used in the proceeding regression analyses.

Table 12. Correlation matrix of IBI component metrics. All component metrics were found to be significantly correlated with overall IBI Score. Components defined in Table 2; HBI and Percent Sensitive values were natural log transformed before analysis.

		TTR	**EPT**	**Becks**	**HBI**	**Shannon**	**% Sensitive**
IBI	r	0.64	0.892	0.919	0.817	0.845	0.873
	p	0.018	0	0	0.001	0	0

Additional Empirical Measures

Water chemical and physical parameters were measured with handheld probes (Table 13). The pH ranged from 5.98 to 7.72, with a mean of 7.27. Dissolved oxygen ranged from 8.72 to 9.92 mg/L, with a mean of 9.36. Conductivity ranged from 52.1 to 223.7 µS, with a mean of 126.9 µS. Specific conductivity ranged from 67.2 to 250.5 µS, with a mean of 154.2 µS. Temperature ranged from 12.8 to 19.4 ^{o}C, with a mean of 15.6 ^{o}C. All measurements were taken between 10:15 a.m. and 1:15 p.m. on 11 June 2015. Data were excluded for Site 14, the most downstream point, because it was originally part of a different study and water chemistry was measured on a separate day in May 2014. Additionally, Site 13 data were excluded from analyses due to the site's ephemeral stream properties.

Grab samples were collected during Summer 2014, Fall 2014, and Spring 2015 (Table 14). Observed nitrate concentrations for the summer samples ranged from 0.46 to 18.81 mg/L, with a mean of 7.03 mg/L. Fall samples ranged from 1.59 to 14.97 mg/L, with a mean of 5.29 mg/L. Spring samples ranged from 0.50 to 17.50 mg/L, with a mean of 5.06 mg/L. The overall averages ranged from 0.48 to 17.09 mg/L, with a mean of 5.57 mg/L. During fall sampling, Site 13 was found to be dry and was thus not sampled. This site drains an upland forested area where nitrate levels appear to be naturally low, so this missing data point helps explain why the fall nitrate values do not drop as low as the summer and spring ranges. Sites 09 and 11, both located

on one tributary, consistently exhibited elevated nitrate levels (Table 14). Although the contributing area to this tributary was chiefly in row crops, which would contribute to the nutrient runoff, the specific reason why this section of the watershed was so highly polluted with nitrate in comparison to other tributaries is unknown.

All three seasonal grab sample values were significantly correlated with the average grab sample values (Table 15). For the additional analyses involving grab samples, the average grab sample was used as the representative variable for all three seasons.

Table 13. Water chemical and physical properties. Site 13 was measured but not included in analyses due to ephemeral stream properties (indicates missing data).*

Site	**pH**	**Dissolved Oxygen (mg/L)**	**Conductivity (µS)**	**Specific Conductivity (µS)**	**Temperature (ºC)**	**Time**
01	7.52	9.27	133.9	157.1	17.3	13:06
02	7.72	9.21	158.1	186.6	17	13:12
03	7.04	9.49	79	97.2	15.2	12:38
04	7.42	9.1	138.1	165.6	16.3	12:30
05	7.87	9.14	223.7	250.5	19.4	12:00
06	7.56	9.72	124.9	154	15.2	11:52
07	6.9	9.17	82.1	101.9	14.85	12:48
08	7.48	8.72	109.4	126.9	18.3	10:20
09	6.9	9.92	191.8	244.6	13.7	10:55
10	7.48	9.49	85.6	105.3	15.2	11:05
11	5.98	9.73	186.4	242.5	12.8	10:42
12	7.29	9.5	84.8	104.8	15	11:15
13	7.35	9.18	52.1	67.2	13.25	12:10
14	*	*	*	*	*	*

Table 14. Grab sample nitrate concentrations (mg/L) for three seasons and the mean.

	Summer	**Fall**	**Spring**	**Average**
Site	**NO_3 (mg/L)**	**NO_3 (mg/L)**	**NO_3 (mg/L)**	**NO_3 (mg/L)**
01	7.07	3.92	3.87	4.96
02	5.56	3.39	2.64	3.86
03	5.28	4.20	3.86	4.45
04	7.95	4.10	4.05	5.37
05	6.98	2.83	5.96	5.26
06	4.09	3.38	1.98	3.15
07	5.74	6.01	4.85	5.53
08	3.77	1.59	2.13	2.50
09	18.71	14.38	16.02	16.37
10	4.02	3.00	2.81	3.28
11	18.81	14.97	17.50	17.09
12	3.74	3.06	2.70	3.17
13	0.46	*	0.50	0.48
14	*	*	3.50	3.50
Outlet	6.24	3.88	3.54	4.55

Table 15. Correlation matrix of average grab sample nitrate with seasonal grab sample nitrate concentrations.

		Summer Nitrate	**Fall Nitrate**	**Spring Nitrate**
Average Nitrate	*r*	0.925	0.786	0.961
	p	0.000	0.001	0.000

Regression

The first iteration of Best Subsets Regression was used to compare models from the five final SWAT variables (Table 4) against the IBI Score. The best two models were generated for each number of variables, including up to five variables. The best equation was determined by the combination of variables that gave a relatively high increase in R^2, R^2 adjusted, and R^2 predictive with their addition to the model, a low Mallows C_p value close to the number of

variables included plus the constant, and a relatively low S value (Table 16). Given these parameters, the best equation was found to include two variables: "Subbasin-Site Only Winter Phosphorus" and "Reach-All Upstream Spring Phosphorus" (R^2=56.0).

Table 16. Best Subsets Regression of the final five SWAT variables on IBI Score. The best equation is highlighted and includes SSWP and RUSpP. Table abbreviations are as follows: SSFaN3, "Subbasin-Site Only Fall Nitrate;" SSWP, "Subbasin-Site Only Winter Phosphorus;" SSSuN3, "Subbasin-Site Only Summer Nitrate;" SSSpN3, "Subbasin-Site Only Spring Nitrate;" and RUSpP, "Reach-All Upstream Spring Phosphorus."

# Var	R^2	R^2 adj	R^2 pred	Mallows C_p	S	SUFaN3	SSWP	SSSuN3	SSSpN3	RUSpP
1	44.0	38.9	27.9	0.1	13.1					X
1	16.8	9.3	0.0	4.5	16.0		X			
2	***56.0***	***47.2***	***33.5***	***0.1***	***12.2***		***X***			***X***
2	47.4	36.8	28.7	1.5	13.3				X	X
3	56.8	42.3	32.4	2.0	12.7		X	X		X
3	56.7	42.2	33.8	2.0	12.7	X	X			X
4	56.9	35.3	25.7	4.0	13.5	X	X	X		X
4	56.8	35.2	17.2	4.0	13.5	X	X		X	X
5	56.9	26.1	3.5	6.0	14.4	X	X	X	X	X

These two SWAT variables were then used in the Fit Regression Model tool (Minitab 17) to generate the regression equation and accompanying statistics. The model was statistically significant (F(2)=6.36, p=0.017) and described by the equation:

$$IBI = 92.2 - 6.76\,SSWP - 13.29\,RUSpP$$

The second iteration of Best Subsets Regression included these two SWAT variables and also added in the three Percent Land Cover variables: "Subbasin-Site Only Percent Forest," "Subbasin-All Upstream Percent Forest," and "Buffer-Site Only Percent Forest." Following the same procedure as above, there were two equally well fitting equations to consider, both with three variables. The first included both of the SWAT variables and "Buffer-Site Only Percent

Forest" (R^2=81.2) and the second included both of the SWAT variables and "Subbasin-Site Only Percent Forest" (R^2=81.0) (Table 17).

Table 17. Second iteration of Best Subsets Regression, including SWAT variables and Percent Forest variables. Table abbreviations are as follows: SSWP, "Subbasin-Site Only Winter Phosphorus;" RUSpP, "Reach-All Upstream Spring Phosphorus;" SSFor, "Subbasin-Site Only Percent Forest;" SUFor, "Subbasin-All Upstream Percent Forest;" and BSFor, "Buffer-Site Only Percent Forest."

# Vars	R^2	R^2 adj	R^2 pred	Mallows C_p	S	SSWP	RUSpP	SSFor	SUFor	BSFor
1	55.8	51.8	33.4	9.3	11.6			X		
1	44.4	39.3	19.7	14.0	13.0					X
2	74.0	68.8	53.7	3.7	9.4		X			X
2	72.7	67.3	51.7	4.3	9.6		X	X		
3	***81.2***	***74.9***	***61.1***	***2.8***	***8.4***	***X***	***X***			***X***
3	***81.0***	***74.7***	***63.7***	***2.8***	***8.4***	***X***	***X***	***X***		
4	82.2	73.3	53.2	4.3	8.7	X	X	X		X
4	82.2	73.3	54.5	4.4	8.7	X	X	X	X	
5	83.0	70.9	40.5	6.0	9.0	X	X	X	X	X

These two models were then used in the Fit Regression Model tool (Minitab 17) to generate the regression equation and accompanying statistics. The first model was statistically significant (F(3)=12.95, p=0.001) and described by the equation:

$$IBI = 73.8 - 5.31\ SSWP - 11.31\ RUSpP + 0.274\ BSFor$$

The second model was also statistically significant (F(3)=12.83, p=0.001) and described by the equation:

$$IBI = 68.6 - 5.67\ SSWP - 9.06\ RUSpP + 0.382\ SSFor$$

The third iteration of Best Subsets Regression included these two SWAT variables and two Percent Land Cover variables, and also added in "SWR Score." There were two equally well fitting equations generated, both with three variables. Both included one SWAT variable,

one Land Cover variable, and SWR score (Table 18). The first equation was slightly better than the second and included "Reach-All Upstream Spring Phosphorus," "Subbasin-Site Only Percent Forest," and "SWR Score" (R^2=86.2). The second equation also included "Reach-All Upstream Spring Phosphorus" and "SWR Score," but used the other land cover variable "Buffer-Site Only Percent Forest" (R^2=84.8).

Table 18. Third iteration of Best Subsets Regression of SWAT, Land Cover, and SWR variables on IBI. Table abbreviations are as follows: SSWP, "Subbasin-Site Only Winter Phosphorus;" RUSpP, "Reach-All Upstream Spring Phosphorus;" SSFor, "Subbasin-Site Only Percent Forest;" SWR, "SWR Score;" and BSFor, "Buffer-Site Only Percent Forest."

Vars	**R^2**	**R^2 adj**	**R^2 pred**	**Mallows C_p**	**S**	**SSWP**	**RUSpP**	**SSFor**	**SWR**	**BSFor**
1	55.8	51.8	33.4	14.9	11.6			X		
1	45.1	40.1	20.4	20.7	13.0				X	
2	74	68.8	53.7	7.1	9.4		X			X
2	72.9	67.5	40.4	7.7	9.5		X		X	
3	***86.2***	***81.6***	***68.9***	***2.5***	***7.2***		***X***	***X***	***X***	
3	***84.8***	***79.8***	***68.2***	***3.2***	***7.5***		***X***		***X***	***X***
4	86.9	80.3	48.2	4.1	7.4	X	X	X	X	
4	86.3	79.5	61.7	4.4	7.6		X	X	X	X
5	87.1	77.8	37.7	6	7.9	X	X	X	X	X

These two models were then used in the Fit Regression Model tool (Minitab 17) to generate the regression equations and accompanying statistics. The first model was statistically significant (F(3)=16.8, p<0.001) and described by the equation:

$$IBI = 28.3 - 10.63\,RUSP + 0.2077\,BSFor + 58.2\,SWR$$

The second model was also statistically significant (F(3)=18.7, p<0.001) and was described by the equation:

$$IBI = 20.6 - 8.75\,RUSP + 0.297\,SSFor + 62.4\,SWR$$

To assess the explanatory value of the grab sample data, the Average Grab Nitrate was added to a Best Subsets Regression on IBI Score with the variables identified as significant in the third iteration: "Reach-All Upstream Spring Phosphorus," "Subbasin-Site Only Percent Forest," "Buffer-Site Only Percent Forest," and "SWR Score." The best models remained the same from the third iteration. Further, Average Grab Nitrate was not selected for any of the one- and two-variable models either (Table 19).

Table 19. Best Subsets Regression of SWAT, Land Cover, SWR, and Grab Sample Nitrate on IBI Score. Table abbreviations are as follows: GAvN3, "Grab Sample Average Nitrate;" RUSpP, "Reach-All Upstream Spring Phosphorus;" SSFor, "Subbasin-Site Only Percent Forest;" SWR, "SWR Score;" and BSFor, "Buffer-Site Only Percent Forest."

# Vars	R^2	R^2 adj	R^2 pred	Mallows Cp	S	GAvN3	RUSpP	SSFor	SWR	BSFor
1	55.8	51.8	33.4	13.6	11.6			X		
1	45.1	40.1	20.4	19.1	13.0				X	
2	74.0	68.8	53.7	6.3	9.4		X			X
2	72.9	67.5	40.4	6.9	9.5		X		X	
3	***86.2***	***81.6***	***68.9***	***2.1***	***7.2***		***X***	***X***	***X***	
3	***84.8***	***79.8***	***68.2***	***2.8***	***7.5***		***X***		***X***	***X***
4	86.3	79.5	61.7	4.0	7.6		X	X	X	X
4	86.2	79.3	60.3	4.1	7.6	X	X	X	X	
5	86.3	76.6	52.8	6.0	8.1	X	X	X	X	X

Finally, to verify that the representative SWAT variable identified in Best Subsets Regression ("Reach-All Upstream Spring Phosphorus") is truly representative of the other variables with which it correlates, the other 12 correlated variables from the first column of Table 9 were regressed in its place with "Subbasin-Site Only Percent Forest" and "SWR Score" on IBI Score. This was to confirm that this variable was not predictive of IBI score by chance alone and that the other variables in its representative group were predictive of IBI score in the same way. Indeed, all 12 other models with the substitute variables were statistically significant

and generated high R^2 values (Table 20). The most significant model was the one using "Reach-All Upstream Summer Organic Nitrogen" ($F(3)=24.9$, $p<0.001$).

Table 20. Comparison of correlated SWAT variables in final model of IBI.

Variable	F	p	R^2	R^2 adjusted	R^2 predictive
Reach-All Upstream Summer Organic Nitrogen	24.9	0.000	89.2	85.7	80.6
Reach-All Upstream Spring Organic Nitrogen	18.8	0.000	86.2	81.6	71.3
Reach-All Upstream Spring Phosphorus	18.7	0.000	86.2	81.6	68.9
Reach-All Upstream Winter Nitrate	17.7	0.000	85.5	80.7	70.5
Reach-All Upstream Annual Organic Nitrogen	17.6	0.000	85.4	80.6	68.8
Reach-All Upstream Annual Phosphorus	15.7	0.001	84.0	78.6	58.5
Reach-All Upstream Winter Phosphorus	14.3	0.001	82.7	76.9	56.5
Reach-All Upstream Annual Nitrate	13.1	0.001	81.4	75.2	57.8
Reach-All Upstream Summer Phosphorus	12.8	0.001	81.0	74.7	46.2
Reach-All Upstream Winter Organic Nitrogen	11.5	0.002	79.3	72.4	53.1
Reach-All Upstream Summer Nitrate	11.4	0.002	79.2	72.3	53.8
Reach-All Upstream Spring Nitrate	11.3	0.002	79.0	72.0	50.0
Buffer-All Upstream Annual Phosphorus	8.7	0.005	74.3	65.7	29.9

DISCUSSION

The purpose of this study was to explore and define the relationship between nutrient runoff and stream ecological condition. SWAT was selected to provide nutrient data at a spatial and temporal resolution that is typically not feasible with empirical measures such as grab sampling. Stream condition was measured using three tiers of ecological assessments: Percent land cover, a Level One GIS-based approach; the SWR Index (Brooks et al. 2009), a Level Two rapid in-field assessment; and the PA-DEP Macroinvertebrate IBI (PA-DEP 2012a), an intensive Level Three in-field and laboratory-based assessment. While the relationship between SWAT-modeled nutrient runoff and macroinvertebrate community integrity was of primary interest, it was important to also include Level One and Level Two approaches. The literature overwhelmingly supports the fact that macroinvertebrates respond to other factors besides nutrients (Richards et al. 1993, Miltner and Rankin 1998, Genito et al. 2002, Riseng et al. 2004, Yuan and Norton 2004, Hering et al. 2006, Kyriakeas and Watzin 2006, Wang 2007, Tran et al. 2010, Wagenhoff et al. 2012). These additional ecological assessments were included to more fully describe how nutrients are linked with macroinvertebrate community condition in the context of other influential factors.

The conclusions from the three ecological approaches were, in general, in agreement and indicated that the majority of the streams in WE38 have been impacted by extensive agricultural activities in the study watershed. Percent land cover metrics identified low percentages of forested area around many study sites. In particular, the “100-m Buffer-Site Only” analysis showed only two sites, Sites 09 and 10, with >80% forested land cover in the buffer around sample reaches. In this same 100-m buffer analysis, five sites were identified as having <20% forested land cover in the buffer (Figure 8). Practically all of the non-forested buffer areas in the

WE38 watershed were found to be used for some aspect of agricultural production, including row crop cultivation and livestock grazing.

The SWR Index also suggested that streams in WE38 have been degraded. The mean SWR score for the watershed was 0.59, indicating a moderate level of impact. One of the component metrics of this multimetric index includes a checklist of stressors affecting the stream, its associated wetlands (if present), and its riparian area 50-m upland in each direction perpendicular to the stream reach. The stressor checklists for the majority of sites indicated that agricultural activities (e.g., row crops, livestock grazing) were directly stressing at least one of these three components.

IBI scores were also indicative of degradation throughout the WE38 watershed, as many sites scored below the cutoff for automatic impairment designation (PA-DEP 2012a). However, the IBI score alone cannot identify causes of degradation, only that impacts to the system exist. Because the Tier 1 and 2 analyses identified activities related to agricultural as likely causes of ecological impairment, it was probable that nutrient runoff would be a contributing stressor to the macroinvertebrate communities. Although the mean condition of streams in WE38 was moderately degraded, there were both higher-integrity and highly degraded sites. Site 05 received an IBI score of <25 and Site 10 received an IBI score of >75, establishing a broad range of scores where the macroinvertebrate-nutrient relationship could be explored.

In prior studies attempting to link SWAT nutrient output and macroinvertebrates, consideration for the scale of summary outputs had been overlooked. The averaging of nutrient values across space and through time holds the potential to dilute key patterns of runoff that impact stream biological integrity. Pionke et al. (2000) found that the majority of annual phosphorous export in the study watershed, WE38, occurred during narrow time windows that

coincided with the seven largest storm events. It was thought that average annual values could not express these high magnitude stressor events as well as a seasonal value. Additionally, because the SWAT model does not route nutrients across the landscape to the stream and instead assumes they reach the stream unprocessed and in their entirety, the spatial scale of summarizing nutrient output was also investigated. It was thought that perhaps the nutrient runoff closest to the stream, in the "Buffer" or at the "Site-Only" scale, would be more closely linked to macroinvertebrates than the whole basin.

There were four research questions guiding this study. First, how could SWAT output be summarized and post-processed to provide an array of nutrient runoff data variables at various spatial and temporal scales that may be valuable to ecologists? Second, how related or unique were each of these nutrient variables, and did the various spatial and temporal scales each convey different information about nutrient runoff in WE38, or were some redundant? Third, what was the relationship between SWAT-modeled nutrients and stream ecological condition, as measured by the macroinvertebrate IBI? And fourth, because macroinvertebrates are known to be impacted by other variables in addition to nutrients, could the relationship between nutrient runoff and stream condition be improved by including additional land use, site condition, or water chemistry variables? Which of these variables were most relevant to explaining macroinvertebrate community integrity?

This study showed that there are many ways to summarize SWAT outputs. In total six spatial scales and five temporal scales were extracted from the model, combining to form 30 unique spatio-temporal scales. The spatial scales were the "Subbasin," "Buffer," and "Reach," each with "Site Only" and "All Upstream" contributing areas. The temporal scales consisted of Annual, Winter, Spring, Summer, and Fall. The "Subbasin-Site Only" and "Reach-All

Upstream" scales were readily available from SWAT output directly; the other four scales were created during post-processing. Three nutrient variables – Total Phosphorus, Total Nitrate, and Total Organic Nitrogen – were selected for 14 sampling sites at each of these 30 spatiotemporal scales. These 90 variables (30 spatiotemporal scales with 3 nutrients each) were then compared to one another to reduce the number of independent variables included in models.

Spearman rank correlations detected high redundancy among SWAT variables. The variables grouped out into five general categories:

1. All "Reach-All Upstream" variables (all seasons, all nutrients)
2. Total phosphorous of "Subbasin-Site Only" and "Buffer-Site Only" scales (all seasons)
3. Total nitrate and organic nitrogen of "Subbasin-Site Only" and "Buffer-Site Only" scales (all seasons except summer)
4. Summer nitrate of "Subbasin-Site Only" and "Buffer-Site Only" scales
5. "Subbasin-All Upstream" and "Buffer-All Upstream" variables (all seasons, all nutrients)

The buffer scale was included in this study to describe the nutrient runoff generated from within 100 m of each stream only. It was correlated in all cases with the subbasin scale variables, and thus, was redundant. I recommend that future studies do not need to include buffer scales in analyses, except in cases where the land use, soil, or topography differs substantially between the buffer and the remainder of the subbasin and where the land area of the subbasin contributing area is much larger than the land area of the buffer contributing area. These two scenarios were not present in WE38, as it is a small watershed that does not have

especially different riparian land characteristics in the buffer zone compared to the rest of the land area.

Another important finding from these groupings is that all of the temporal scales were generally redundant. Annual and seasonal variables were all correlated, with the exception of summer at two spatial scales for nitrate. The same pattern of nutrient runoff was observed at all sites regardless of the time of year, though not necessarily the same in magnitude. This held true for both SWAT-modeled nutrient and grab sample nitrate concentrations. Given the importance of seasonal runoff events, however, it is recommended that future studies consider including seasonal variation to verify these observations.

Lastly, it was seen that phosphorous, nitrate, and organic nitrogen tended to follow the same patterns within each spatial scale. Phosphorous did diverge from nitrate and organic nitrogen in the "Subbasin-Site Only" and "Buffer-Site Only" scales. Deciding to choose one representative nutrient in future studies may be dependent on the specific nutrient sources in the watershed of interest. If phosphorous and nitrogen sources overlap, then including both nutrients may not be necessary.

Best Subsets Regression using five representative SWAT variables from the five unique groupings showed a significant negative relationship between instream nutrient concentration and IBI score. The best model included "Subbasin-Site Only Winter Phosphorus" and "Reach-All Upstream Spring Phosphorus." This reach variable represented all of the "Reach-All Upstream" nutrients. Its inclusion in the model pointed to the importance of nutrient runoff from both the upstream contributing areas and the smaller site subbasin. The exclusion of "Subbasin-All Upstream" from the model pointed to the importance of instream processing and other mathematical functions that the Reach output included; the Reach included the same nutrient

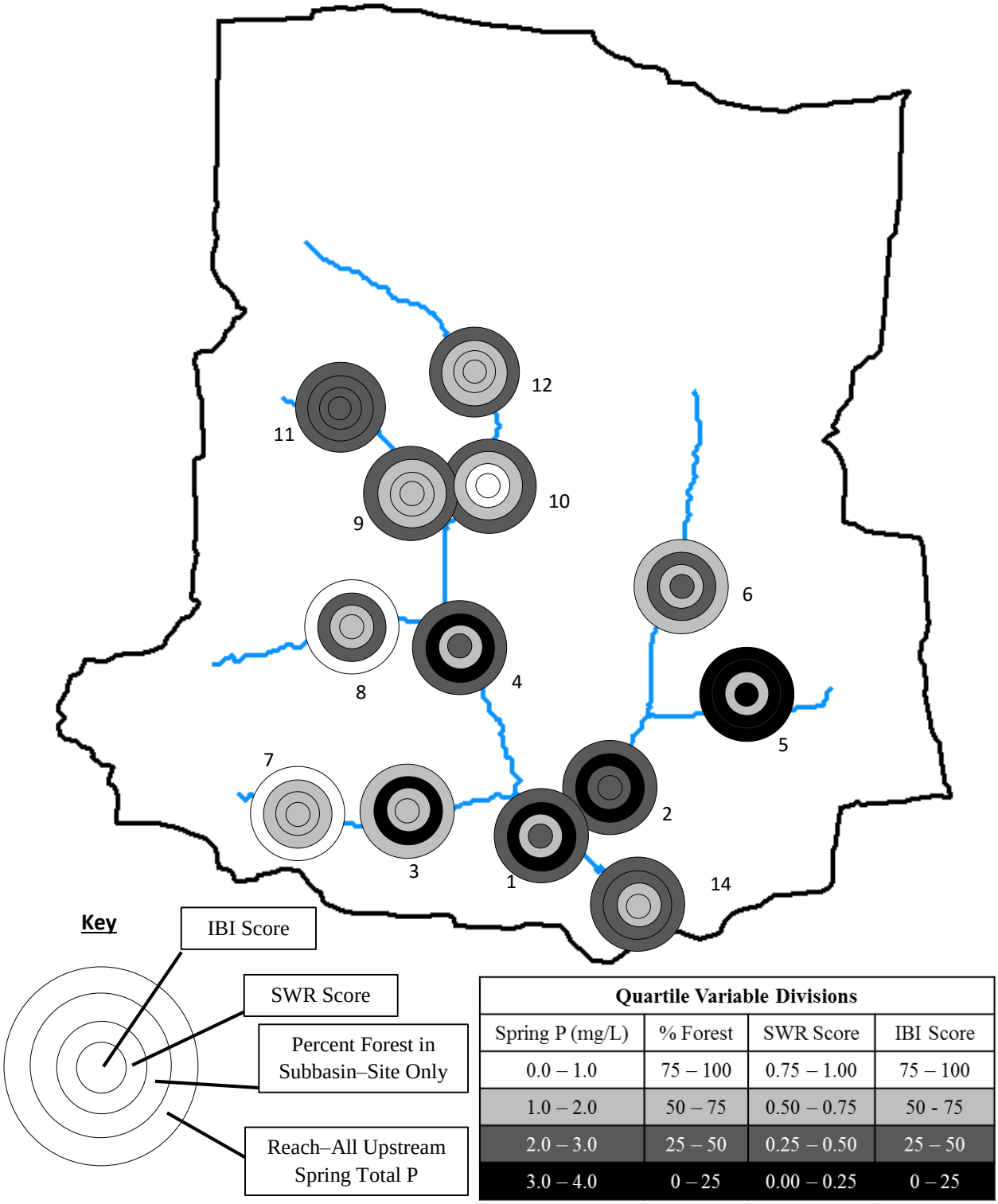
data as Subbasin, but also calculated instream processing from upstream contributing reaches. Lastly the inclusion of the "Subbasin-Site Only" variable showed that site-level nutrient runoff specifically was also an important factor.

During the final regressions that included non-SWAT variables, three variables were identified in the best model: "Reach-All Upstream Spring Phosphorous," "Subbasin-Site Only Percent Forest," and SWR score. As expected, the regression model was improved by including these additional ecological variables, increasing R^2 from 56.0 to to 86.2. This improvement in regression fit demonstrates the direct impact potential of land use surrounding sampling areas on macroinvertebrate community integrity. Higher forest coverage can relate to higher canopy shading, bank and soil stability, and pollutant filtration. Canopy shading helps to dampen wide fluctuations in temperature that can be stressful to aquatic organisms. Bank and soil stability are important for decreasing sediment runoff potential and instream bank erosion, which can decrease interstitial habitat space and act as physical stressors to macroinvertebrate respiration. Pollutant filtration is another benefit of forested riparian buffers, intercepting and processing pollutants before they pollute adjacent water bodies (Karr and Schlosser 1978).

The contribution of the SWR score to explaining macroinvertebrate community was also of note. This index was developed by Brooks et al. (2009) to identify the factors in the stream, wetland, and riparian portions of a sample site that are most indicative of biological integrity. This index provided an overview of the environmental conditions that macroinvertebrates were exposed to in greater detail than the landscape level could provide. The single overall score could then be used in regression models, eliminating the need to separately include variables such as stream incision, instream habitat quality, and contributing stressors. This reduction of factors reduces model overfitting and produces a stronger model, as demonstrated in this study.

Somewhat surprisingly, grab sample concentration of nitrate was not included as a variable in the final regression models. Although SWAT-modeled nutrients were found to be important explanatory factors, empirically collected nitrate was not. This was presumed to be because of the scarcity of grab sampling, which occurred once during the spring, summer, and fall seasons at each site. The sporadic sampling design provided only a snapshot of fluctuating nutrient conditions to which macroinvertebrates were exposed, supporting the idea that infrequent grab samples are unreliable in representing overall nutrient runoff levels. It may also explain why Genito et al. (2002) did not find the same strong relationship between macroinvertebrates and nutrients in the WE38 watershed. Other studies have experienced similar issues, concluding that no relationship was found between macroinvertebrates and nutrients when only grab samples were available (e.g., Richards et al. 1993, Calhoun et al. 2008).

The interplay between the three variables selected in the best model was interesting in context of the specific sites sampled in WE38. Figure 11 shows how each site scored for each of the three predictor variables as well as its IBI score, represented in quartiles for simplicity. The outermost ring shows SWAT-modeled "Reach-All Upstream Spring Phosphorous." The second ring in shows the "Subbasin-Site Only Percent Forest." The third ring shows "SWR Score" and the innermost circle shows "IBI Score." Darker colors relate to more degraded conditions (i.e., higher nutrients and lower Percent Forest, SWR, and IBI). Lighter colors relate to healthier conditions (i.e., lower nutrients and higher Percent Forest, SWR, and IBI).



Quartile Variable Divisions			
Spring P (mg/L)	% Forest	SWR Score	IBI Score
0.0 – 1.0	75 – 100	0.75 – 1.00	75 – 100
1.0 – 2.0	50 – 75	0.50 – 0.75	50 - 75
2.0 – 3.0	25 – 50	0.25 – 0.50	25 – 50
3.0 – 4.0	0 – 25	0.00 – 0.25	0 – 25

Figure 11. Nutrient level, Percent Forest, SWR Score, and IBI Score for each sample site in WE38. Each concentric circle represents a different variable, and the color represents the value of that variable. Darker colors relate to higher degradation, while lighter colors represent healthier conditions.

In general, these metrics tended to score similarly, such as in Site 11, where all three predictor variables were in the third quartile and the IBI score was also in the same. More interestingly, perhaps, were the sites where predictor metrics were contradictory. Site 10 had moderately poor nutrient runoff that was likely due to the intensive agricultural activities upstream, where the riparian trees and shrubs had been removed and livestock had full access to the stream. The Percentage Forest at the "Subbasin-Site Only" scale was fair due to a large forested patch among the agricultural fields. The SWR Score was optimal, due in part to a forested buffer, high habitat quality, low incision ratio, and intact floodplain wetlands. The IBI Score was also optimal. The macroinvertebrates identified from Site 10 included genera with pollution tolerance values of 0, the most pollution-sensitive taxa, including *Diplectrona, Epeorus, Haploperla, and Leuctra spp.* This site was exemplary of how preservation of the adjacent riparian zone and associated wetlands and other site-level factors can counteract the negative impacts of upstream agriculture and runoff, providing a rescue effect on the ecological community within that high-integrity site.

To the contrary, Site 5 exemplified how site condition can only do so much for the ecological community when paired with high nutrient runoff and high deforestation in the subbasin-site scale (e.g., conversion to agriculture, as was the case here). This site had the highest modeled nutrient runoff and second lowest percent forest cover in the entire watershed. Its SWR score, however, was only suboptimal. Despite a relatively higher score for the site condition, the IBI score was also the lowest out of the entire watershed. The macroinvertebrate community was dominated by chironomidae individuals, a moderately pollution tolerant family of true fly larvae, that made up 73.3% of the macroinvertebrate sample. A potato farm at the headwaters of this stream is likely to be the major cause of the degradation. It was observed that

farther upstream of Site 5, the channel was deeply incised and banks were unstable. Despite the channel geomorphology shifting to a more stable condition at the sample site and some riparian trees and shrubs growing, the nutrient and land use changes upstream outweighed these attributes.

The patterns observed at sites within WE38 generally agree with other studies that describe a wedge-shaped response of macroinvertebrates to nutrients (e.g., Figure 2). Wang et al. (2007) explained this response in that, at low nutrient concentrations, other variables are more impactful on stream ecology. At high nutrient concentrations, however, the response is almost always impacted. This agrees with the observations at Sites 10 and 5 in particular. These patterns also highlight the importance of improving both the physical habitat and the nutrient pollution of impaired streams. The habitat effect seen at Site 10 agrees with the findings of Moore and Palmer (2005) that maintenance of intact riparian buffers can help mitigate the effects of nutrient pollution from agriculture. The nutrient reduction required at Site 5 to see any benefit from its higher quality habitat also supports the nationwide efforts to find solutions to the systemic nutrient runoff impacting our waterways.

This study has added to the growing body of literature addressing nutrient impacts on stream ecology, providing further evidence that there is a group of nutrient runoff variables within the SWAT model output that are tightly related to stream biological integrity as measured by IBI Score. The issue of scale was addressed, as SWAT outputs showed the “Reach-All Upstream” spatial scale to be most relevant and land use variables showed that the “Subbasin-Site Only” scale was most important. Woznicki et al. (2015) also studied the relationship among SWAT nutrient variables and macroinvertebrate community metrics, and found nutrient relationships with the “Reach-All Upstream” scale, although additional spatial scales were not

included. Findings from other studies that used empirical measures of nutrients were also considered. Sponseller et al. (2001) concluded that macroinvertebrates were most related to site conditions at the 200-m subcorridor scale, but were not related to water chemistry, which was more related to catchment-scale conditions. Calhoun et al. (2008) concluded that macroinvertebrates were related to the agricultural land use near streams, but not the nutrient water chemistry. These studies perhaps show incomplete conclusions due to the limitations of sparse, empirically collected nutrient data. They do agree with the finding in this study that land use in the area immediately surrounding the stream is the most influential in shaping the macroinvertebrate community. Tran et al. (2010) also found that the 200-m buffer zone was most impactful on macroinvertebrates, though nutrients were not included in the study.

It is vital to thoroughly explore SWAT and define the variables most relevant to ecological condition because there has been a clear interest at the national level in using SWAT for ecological applications, as evidenced by substantial grant projects currently funded by federal organizations. The USEPA currently has its STAR grant program funding the Center for Integrated Multi-Scale Nutrient Pollution Solutions (http://agsci.psu.edu/aec/research-extension/research-centers/center-for-nutrient-solutions) and the USDA NRCS currently has its CEAP grant program funding the Great Lakes CEAP (http://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/technical/nra/?cid=stelprdb1045403). Assuming the investment in these types of applied, cross-disciplinary opportunities with SWAT continues, it is judicious to set up the connections between model outputs and ecological condition as soon as possible. This project adds to the small but rapidly growing literature on the subject.

The “Reach-All Upstream” spatial scale should be included for any future studies seeking to link macroinvertebrates with SWAT-modeled nutrient data. Additionally, this study supports the inclusion of local environmental conditions that surround macroinvertebrate sampling reaches in nutrient-related studies. By accounting for factors external to nutrient runoff, the relationship between macroinvertebrates and nutrient pollution can be better understood.

While this study attempted to include a variety of parameters thought to be important in shaping macroinvertebrate communities, it is likely that other important factors were overlooked. Woznicki et al. (2016) transferred output from the SWAT model to the Hydrologic Index Tool (HIT) to generate additional variables on hydrology and flow for sample sites. They found that these variables, especially the duration and timing of low flow events, were highly influential in explaining macroinvertebrate community composition. Future studies should consider including these additional factors in analyses.

Another limitation to this study is the reliance on and assumption of linear relationships. Specifically, linearity may not be the most appropriate assumption for the regression analyses. While linear models are simple and can be effective, ecological responses are often more complex. Einheuser et al. (2012, 2103) found that macroinvertebrate and fish indicators are better suited to nonlinear modeling, such as the adaptive-neuro fuzzy inference systems (ANFIS) used in their studies. Nonlinear approaches, however, were beyond the scope of this thesis. The use of Spearman rank correlation as a variable reduction technique was recommended by Einheuser et al. (2012).

One major question remains for anyone using the SWAT model in place of empirical measurements of nutrients: how valid are modeled nutrient values? To truly validate the model for specific watershed, researchers would need the high-resolution nutrient data and land

management data that this approach is designed to avoid. There is an inherent leap of faith that the model is a relatively accurate representation of real-world runoff patterns through space and time. One way to increase the confidence of SWAT outputs is to use the best input data possible. In this study, the WE38 watershed has been studied by the USDA-ARS for over three decades and was modeled using an extreme amount of historical management and experimental data (Veith et al. 2015). It is a small area to model, which translates to a large amount of data per unit of land area. The relationships described in this watershed among environmental variables and stream integrity are strong and clear, but this clarity may diminish as modeled watersheds get larger, data quality diminishes, and natural system variations increase.

Further work is needed to test the findings of this small-scale study in east-central Pennsylvania and to verify that the Reach-All Upstream spatial scale remains important in predicting IBI Score for other watersheds and in other physiographic regions throughout the Chesapeake Bay watershed and beyond. This study was conducted in a small, data-rich watershed where landowner compliance and participation was exemplary. The modeling of WE38 by the USDA-ARS Pasture and Watershed Management Research Unit has been ongoing for years and has continued to be refined and improved to provide some of the best SWAT modeling possible for a study on this topic. Acknowledging that the efforts of the USDA-ARS and related researchers in WE38 are atypical of most watersheds in the Chesapeake Bay watershed, this research should be replicated at larger scales and within areas where data and prior modeling work is less available.

Given the tradeoffs involved with management decisions like fertilizer application, land development, and implementation of best management practices (BMPs), the nutrient challenge is not an easy dilemma to overcome. The range of scales involved with Chesapeake Bay issues

is one reason for this complexity, as decisions in the upstream jurisdictions (e.g., Pennsylvania or New York) impact the downstream resources in the Bay. There are local ecosystem services in headwater regions, but most discourse has centered on improving Bay condition. How do we protect and sustain ecosystems that are spatially remote from stressor inputs? How do we encourage sustainable practices upstream of major visible impacts? How do we provide the material, food, and energy required by society within the constraints of legislation that supervises production systems? There are no simple solutions to these questions. Fortunately, research can help break down the false dichotomy of human resource consumption versus environmental sustainability and find solutions that work for upstream and downstream systems alike.

As nutrient models such as SWAT are developed for more watersheds and at larger spatial scales, it is judicious to begin setting up the connections between model outputs and ecological condition. Macroinvertebrates are the most commonly used group for biological monitoring, so this particular connection is especially important. Various recommendations on how to overcome the nonpoint source pollution problem include increasing fertilizer use efficiency, maximizing denitrification, reducing runoff from agricultural and urban areas, and linking living resources to eutrophication (Carpenter et al. 1998, Boesch et al. 2001, Schlesinger 2009). All of these would benefit from a spatially explicit view of nutrient source areas to make interventions as effective as possible. Additionally, the ability to identify any local benefits that landowners in headwater areas might experience may provide vital incentives for them to implement Best Management Practices (BMPs) on their lands. The propensity to monitor ecological condition within this local nutrient source framework could ultimately be valuable for fulfilling the Clean Water Act's call for "chemical, physical, and biological integrity."

LITERATURE CITED

Ashton, M. J., R. P. Morgan Ii, and S. Stranko. 2014. Relations between Macroinvertebrates, Nutrients, and Water Quality Criteria in Wadeable Streams of Maryland, USA. *Environmental Monitoring and Assessment* 186(2): 1167–82.

Barbour, M..T., J. Gerritsen, B..D. Snyder, and J..B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C. http://water.epa.gov/scitech/monitoring/rsl/bioassessment/

Beven, K. J. and M. J. Kirkby. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin* 24(1): 43-69.

Boesch, D. F., R. B. Brinsfield, and R. E. Magnien. 2001. Chesapeake Bay Eutrophication. *Journal of Environment Quality* 30(2): 303–320.

Boyer, E. W., C. L. Goodale, N. A. Jaworski, and R. W. Howarth. 2002. Anthropogenic Nitrogen Sources and Relationships to Riverine Nitrogen Export in the Northeastern U.S.A. *Biogeochemistry* 57/58: 137–169.

Boynton, W. R., J. H. Garber, R. Summers, and W. M. Kemp. 1995. Inputs, transformations, and transport of nitrogen and phosphorus in Chesapeake Bay and selected tributaries. *Estuaries* 18(1): 285–314.

Bricker, S., B. Longstaff, W. Dennison, A. Jones, K. Boicourt, C. Wicks, and J. Woerner. 2007. Effects of Nutrient Enrichment In the Nation's Estuaries: A Decade of Change. NOAA Coastal Ocean Program Decision Analysis Series No. 26. National Centers for Coastal Ocean Science, Silver Spring, MD. 328 pp.

Brooks, R. P., D. H. Wardrop, and C. A. Cole. 2006. Inventorying and monitoring wetland condition and restoration potential on a watershed basis with examples from the Spring Creek watershed, Pennsylvania, USA. *Environmental Management* 38: 673-687.

Brooks, R., M. McKenney-Easterling, M. Brinson, R. Rheinhardt, K. Havens, D. O'Brien, J. Bishop, J. Rubbo, B. Armstrong, and J. Hite. 2009. A Stream-Wetland-Riparian (SWR) Index for Assessing Condition of Aquatic Ecosystems in Small Watersheds along the Atlantic Slope of the Eastern U.S. *Environmental Monitoring and Assessment* 150: 101-117.

Bryant, R. B., T. L. Veith, G. W. Feyereisen, A. R. Buda, C. D. Church, G. J. Folmar, J. P. Schmidt, C. J. Dell, and P. J. A. Kleinman. 2011, USDA-ARS Mahantango Creek Watershed, Pennsylvania, United States: Physiography and history. *Water Resources Research* 47: W08701. Doi 10.1029/2010wr010056.

Burton, J. and J. Gerritsen. 2003. A Stream Condition Index for Virginia Non-Coastal Streams. Tetra Tech, Inc., Owings Mills, MD. 163 pages. <www.deq.virginia.gov/Portals/0/DEQ/Water/WaterQualityMonitoring/BiologicalMonitoring/vsci.pdf>

Calhoun, Daniel L., M. Brian Gregory, and Holly S. Weyers. 2008. Algal and Invertebrate Community Composition along Agricultural Gradients: A Comparative Study from Two Regions of the Eastern United States. USGS Numbered Series. Scientific Investigations Report. Geological Survey. <http://pubs.er.usgs.gov/publication/sir20085046>

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W Howarth, A. N. Sharpley, and V. H. Smith. 1998. Nonpoint Pollution of Surface Waters with Phosphorus and Nitrogen. *Ecological Applications* 8(3): 559–568.

Chambers, P. A., D. J. McGoldrick, R. B. Brua, C. Vis, J. M. Culp, and G. A. Benoy. 2012. Development of Environmental Thresholds for Nitrogen and Phosphorus in Streams. *Journal of Environmental Quality* 41: 7–20.

Chesapeake Bay Program (CBP). 1983. 1983 Chesapeake Bay Agreement. Annapolis, MD. http://www.chesapeakebay.net/content/publications/cbp_12512.pdf

Chesapeake Executive Council (CEC). 1987. Chesapeake Bay Agreement. Chesapeake Bay Program, Annapolis, Maryland.

Chesapeake Executive Council (CEC). 2000. Chesapeake 2000. Chesapeake Bay Program, Annapolis, Maryland.

Clean Water Act of 1972, 33 U.S.C.§1251 et seq. 2002. <http://epw.senate.gov/water.pdf>

Collick, A. S., D. R. Fuka, P. J. Kleinman, A. R. Buda, J. L. Weld, M. J. White, T. L. Veith, R. B. Bryant, C. H. Bolster, and Z. M. Easton. 2014. Predicting phosphorus dynamics in complex terrains using a variable source area hydrology model. *Hydrological Processes* 29(4): 588-601.

Committee on Environment and Natural Resources (CENR). 2010. Scientific Assessment of Hypoxia in U.S. Coastal Waters. Interagency Working Group on Harmful Algal Blooms, Hypoxia, and Human Health of the Joint Subcommittee on Ocean Science and Technology. Washington, DC.

Correll, D. L. 1998. The Role of Phosphorus in the Eutrophication of Receiving Waters: A Review. *Journal of Environmental Quality* 27: 261-266.

Daggupati, P., H. Yen, M. J. White, R. Srinivasan, J. G. Arnold, C. S. Keitzer, and S. P. Sowa. 2015. Impact of Model Development, Calibration and Validation Decisions on Hydrological Simulations in West Lake Erie Basin. *Hydrological Processes* 29 (26): 5307–20.

Diaz, R. J., and R. Rosenberg. 2008. Spreading Dead Zones and Consequences for Marine Ecosystems. *Science* 321(5891): 926–929.

Dubrovsky, N. M., and P. A. Hamilton. 2010. Nutrients in the Nation's streams and groundwater: National Findings and Implications: U.S. Geological Survey Fact Sheet 2010-3078, 6 pages.

Easton, Z. M., D. R. Fuka, M. T. Walter, D. M. Cowan, E. M. Schneiderman, and T. S. Steenhuis. 2008. Re-Conceptualizing the Soil and Water Assessment Tool (SWAT) model to predict runoff from variable source areas. *Journal of Hydrology* 348: 279-291.

Einheuser, Matthew D., A. Pouyan Nejadhashemi, Scott P. Sowa, Lizhu Wang, Yaseen A. Hamaamin, and Sean A. Woznicki. 2012. Modeling the Effects of Conservation Practices on Stream Health. *Science of The Total Environment* 435(36): 380–91.

Einheuser, Matthew D., A. Pouyan Nejadhashemi, Lizhu Wang, Scott P. Sowa, and Sean A. Woznicki. 2013a. Linking Biological Integrity and Watershed Models to Assess the Impacts of Historical Land Use and Climate Changes on Stream Health. *Environmental Management* 51(6): 1147–63.

Einheuser, Matthew D., A. Pouyan Nejadhashemi, and Sean A. Woznicki. 2013b. Simulating Stream Health Sensitivity to Landscape Changes due to Bioenergy Crops Expansion. *Biomass and Bioenergy* 58: 198–209.

Environmental Systems Research Institute (ESRI). 2012. ArcGIS Desktop: Release 10.1. Redlands, CA. http://www.esri.com

Executive Order No. 13508, 3 C.F.R. 26226 -26227. 2009. <https://www.federalregister.gov/articles/2010/05/11/2010-11143/executive-order-13508-chesapeake-bay-protection-and-restoration-section-203-final-coordinated>

Federal Leadership Committee for the Chesapeake Bay (FLC). 2014. Combined FY 2014 Action Plan and FY 2013 Progress Report: Public Comment Draft. <http://executiveorder.chesapeakebay.net/file.axd?file=2014%2f3%2f2013_Progress+_2014+Action.pdf>

Fuka, D.R., A. S. Collick, P. J. Kleinman, A. Auerbach, R. D. Harmel, and Z. M. Easton. 2015. Improving the spatial representation of soil properties and hydrology using topographically derived initialization processes in the SWAT model. *Hydrologic Processes* (in review).

Genito, D., W. J. Gburek, and A. N. Sharpley. 2002. Response of Stream Macroinvertebrates to Agricultural Land Cover in a Small Watershed. *Journal of Freshwater Ecology* 17(1): 109–19.

Harbridge, J. 2007. Determination of ammonia by flow injection analysis (high throughput) (salicylate method). QuikChem® Method 10-107-06-2-L. Loveland, Colorado: Lachat Insturments.

Harbridge, J. and S. Tucker. 2008. Determination of nitrate/nitrite in surface and wastewaters by flow injection analysis (high throughput). QuikChem® Method 10-107-04-1-R. Loveland, Colorado: Lachat Instruments.

Hering, D., R. K. Johnson, S. Kramm, S. Schmutz, K. Szoszkiewicz, and P. F. M. Verdonschot. 2006. Assessment of European Streams with Diatoms, Macrophytes, Macroinvertebrates and Fish: A Comparative Metric-Based Analysis of Organism Response to Stress. *Freshwater Biology* 51 (9): 1757–85.

Herman, M. R., A. P. Nejadhashemi, F. Daneshvar, D. M. Ross, S. A. Woznicki, Z. Zhang, and A. Esfahanian. 2015. Optimization of Conservation Practice Implementation Strategies in the Context of Stream Health. *Ecological Engineering* 84: 1–12.

Homer, C., J. Dewitz, L. Yang, S. Jin, P. Danielson, G. Xian, J. Coulston, N. Herold, J. Wickham, and K. Megown. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, v. 81, no. 5, p. 345-354.

Jaworski, N. 1990. Retrospective of the water quality issues of the upper Potomac estuary. *Aquatic Science* 3:11–40.

Jordan, T. E., D. L. Correll, and D. E. Weller. 1997. Effects of Agriculture on Discharges of Nutrients from Coastal Plain Watersheds of Chesapeake Bay. *Journal of Environment Quality* 26(3): 836.

Justus B. G., James C. Petersen, Suzanne R. Femmer, Jerri V. Davis, and J. E. Wallace. 2010. A Comparison of Algal, Macroinvertebrate, and Fish Assemblage Indices for Assessing Low-Level Nutrient Enrichment in Wadeable Ozark Streams. *Ecological Indicators* 10(3): 627–38.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. *Fisheries* 6(6): 21-27.

Karr, J. R. 1991. Biological integrity: A long-neglected aspect of water resource management. *Ecological Applications* 1:66-84.

Karr, J. R., and E. W. Chu. 2000. Sustaining Living Rivers. *Hydrobiologia* 422(23): 1–14.

Karr, J. R., and D. R. Dudley. 1981. Ecological perspective on water quality goals. *Environmental Management* 5(1): 55-68.

Karr, J. R., & Schlosser, I. J. (1978). Water resources and the land-water interface. *Science* 201(4352): 229-234.

Kautza, Adam, S. Mazeika, and P. Sullivan. 2012. Using a process-based catchment-scale model for enhancing field-based stream assessments and predicting stream fish assemblages. *Aquatic Conservation: Marine and Freshwater Ecosystems* 22(4): 511–25.

Kemp, W. M., W. R. Boynton, J. E. Adolf, D. F. Boesch, W. C. Boicourt, G. Brush, J. C. Cornwell, et al. 2005. Eutrophication of Chesapeake Bay: Historical Trends and Ecological Interactions." *Marine Ecology Progress Series* 303: 1–29.

Klauda, R., P. Kazyak, S. Stranko, M. Southerland, N. Roth, and J. Chaillou. 1998. Maryland Biological Stream Survey: A state agency program to assess the impact of anthropogenic stresses on stream habitat quality and biota. *Environmental Monitoring and Assessment* 51(2): 299 – 316.

Knobeloch, L., B. Salna, A. Hogan, J. Postle, and H. Anderson. 2000. Blue babies and nitrate-contaminated well water. *Environmental Health Perspective* 108(7): 675–678.

Kyriakeas, S. A., and M. C. Watzin. 2006. Effects of Adjacent Agricultural Activities and Watershed Characteristics on Stream Macroinvertebrate Communities. *Journal of the American Water Resources Association* 42(2): 425-441.

Langland, M.J., J. P. Raffensperger, D. L. Moyer, J. M. Landwehr, G. E. Schwarz. 2006. Changes in Streamflow and Water Quality in Selected Nontidal Basins in the Chesapeake Bay Watershed, 1985-2004: U.S. Geological Survey Scientific Investigations Report 2006-5178. 75 pages.

Linker, L. C., R. A. Batiuk, G. W. Shenk, and C. F. Cerco. 2013. Development of the Chesapeake Bay Watershed Total Maximum Daily Load Allocation. *Journal of the American Water Resources Association* 49(5): 986-1006.

Lyerly, C.M., A.L. Hernández Cordero, K.L. Foreman, S.W. Phillips, W.C. Dennison. 2013. Lessons from Chesapeake Bay Restoration Efforts: Understanding the role of nutrient reduction activities in improving water quality. <http://www.chesapeakebay.net/channel_files/18860/bill_dennison_-_lessons_from_chesapeake_bay_restoration_efforts_-_20131120.pdf>

Merritt, R.W., K. W. Cummins, and M. B. Berg. 2008. An introduction to the aquatic insects of North America. 4th Edition. Dubuque, Iowa: Kendall Hunt Publishing.

Miltner, R. J., and E. T. Rankin. 1998. Primary Nutrients and the Biotic Integrity of Rivers and Streams. *Freshwater Biology* 40(1): 145–58.

Minitab 17 Statistical Software. 2010. Computer software. State College, PA: Minitab, Inc. www.minitab.com

Moore, A. A., and M. A. Palmer. 2005. Invertebrate Biodiversity in Agricultural and Urban Headwater Streams: Implications for Conservation and Management. *Ecological Applications* 15(4): 1169–1177.

Mulholland, P. J., A. M. Helton, G. C. Poole, R. O. Hall, S. K. Hamilton, B. J Peterson, and S. M. Thomas. 2008. Stream denitrification across biomes and its response to anthropogenic nitrate loading. *Nature* 452(7184): 202–205.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, J. R. Williams, and K. W. King. 2002. Soil and Water Assessment Tool: Theoretical documentation. Report TR-191. College Station, Tex.: Texas Water Resources Institute.

New York State Department of Environmental Conservation (NYS-DEC) Division of Water. 2012. Standard operating procedure: Biological monitoring of surface waters in New York state. 164 pages. < www.dec.ny.gov/docs/water_pdf/sop20814final.pdf>

PA Department of Environmental Protection (PA-DEP). 2009. Survey Protocol – Instream Comprehensive Evaluation Surveys. Harrisburg, PA. <http://www.portal.state.pa.us/portal/server.pt/document/F1619621/ice_pdf>

PA Department of Environmental Protection (PA-DEP). 2012a. A benthic index of biotic integrity for wadeable freestone riffle-run streams in Pennsylvania. Division of Water Quality Assessment and Standards, Pennsylvania Department of Environmental Protection, Harrisburg, PA. <http://www.portal.state.pa.us/portal/server.pt/document/1433836/freestoneibimarch2012_pdf_%282%29>

PA Department of Environmental Protection (PA-DEP). 2012b. Pennsylvania's Surface Water Quality Monitoring Network (WQN). 3800-BK-DEP0636. <http://www.dep.pa.gov/Business/Water/PointNonPointMgmt/WaterQuality/Pages/Integrated-Water-Quality-Report-2014.aspx#.VuQyiND1rM4>

Peckarsky, B. L. 1990. Freshwater Macroinvertebrates of Northeastern North America. Ithaca, New York: Comstock Publishing Associates.

Pionke, H. B., B. J. Gburek, and A. N. Sharpley. 2000. Critical Source Area Controls on Water Quality in an Agricultural Watershed Located in the Chesapeake Basin. *Ecological Engineering* 14(4): 325–35.

Richards, C., G. E. Host, and J. W. Arthur. 1993. Identification of Predominant Environmental Factors Structuring Stream Macroinvertebrate Communities within a Large Agricultural Catchment. *Freshwater Biology* 29(2): 285–94.

Riseng, C. M., M. J. Wiley, and R. J. Stevenson. 2004. Hydrologic Disturbance and Nutrient Effects on Benthic Community Structure in Midwestern US Streams: A Covariance Structure Analysis. *Journal of the North American Benthological Society* 23(2): 309–26.

Rosenberg, D. M., and V. H. Resh. 1993. Freshwater biomonitoring and benthic macroinvertebrates. New York: Chapman & Hall.

Schlesinger, W. H. 2009. On the fate of anthropogenic nitrogen. *Proceedings of the National Academy of Sciences* 106(1): 203–208.

Sponseller, R. A., E. F. Benfield, and H. M. Valett. 2001. Relationships between Land Use, Spatial Scale and Stream Macroinvertebrate Communities. *Freshwater Biology* 46(10): 1409–24.

Stewart, K. W., B. P. Stark, and J. A. Stanger. 2002. Nymphs of North American Stonefly Genera (Plecoptera). Second Edition. Thomas Say Foundation (Series), v. 12. Columbus, Ohio: Caddis Press.

Strahler, A. N. 1952. Hypsometric (area-latitude) analysis of erosional topography. *Bulletin of the Geological Society of America* 63: 1117–1142.

Tran, C. P., R. W. Bode, A. J. Smith, and G. S. Kleppel. 2010. Land-Use Proximity as a Basis for Assessing Stream Water Quality in New York State (USA). *Ecological Indicators* 10(3): 727–33.

Trimble Navigation, Limited. 2015. GPS Pathfinder Pro XRS submeter handheld unit, with onsite corrections via OmniSTAR subscription. <http://www.trimble.com> and <http://www.omnistar.com>

U.S. Department of Agriculture – Agricultural Research Service (USDA-ARS). 2011. Mahantango Creek Watershed, Pennsylvania: An ARS Benchmark Research Watershed. Last modified 14 Sep 2011. http://www.ars.usda.gov/Research/docs.htm?docid=18639

U.S. Environmental Protection Agency (USEPA). 1999. Draft Guidance for Water Quality-based Decisions: The TMDL Process (Second Edition). Office of Water, Washington, D.C. http://www.nepis.epa.gov/Exe/ZyPDF.cgi/P1007N47.PDF?Dockey=P1007N47.PDF

U.S. Environmental Protection Agency (USEPA). 2003. Setting and Allocating the Chesapeake Bay Basin Nutrient and Sediment Loads, the Collaborative Process, Technical Tools and Innovative Approaches. Region 3 Chesapeake Bay Program Office, Annapolis, Maryland.

U.S. Environmental Protection Agency (USEPA). 2007. Memorandum: Nutrient Pollution and Numeric Water Quality Standards. From: Benjamin H. Grumbles, Assistant Administrator, U.S. Environmental Protection Agency Office of Water. <www.epa.gov/waterscience/criteria/nutrient/files/policy20070525.pdf>

U.S. Environmental Protection Agency (USEPA). 2010a. Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorus and Sediment. U.S. Environmental Protection Agency Chesapeake Bay Program Office, Annapolis, Maryland. http://www.epa.gov/reg3wapd/tmdl/ChesapeakeBay/tmdlexec.html

U.S. Environmental Protection Agency (USEPA). 2010b. Guide for EPA's Evaluation of Phase I Watershed Implementation Plans. U.S. Environmental Protection Agency, Region 3, Philadelphia, Pennsylvania. http://www.epa.gov/reg3wapd/pdf/pdf_chesbay/GuideforEPAWIPEvaluation4-2-10.pdf

U.S. Environmental Protection Agency Science Advisory Board (USEPA SAB). 2011. Reactive Nitrogen in the United States: An analysis of Inputs, Flows, Consequences, and Management Options. EPA-SAB-11-013.

U.S. Geological Survey (USGS). 1999. Monitoring Nutrients in the Major Rivers Draining to Chesapeake Bay. Water-Resources Investigations Report 99-4238.

Veith, T. L., J. E. Richards, S. C. Goslee, A. S. Collick, R. B. Bryant, D. A. Miller, B. W. Bills, A. R. Buda, R. L. Sebring, and P. J. A. Kleinman. 2015. Navigating spatial and temporal complexity in developing a long-term land use database for an agricultural watershed. *Journal of Soil and Water Conservation* 70: 288-96.

Vitousek, P. M., J. D. Aber, R. W. Howarth, G. E. Likens, P. A. Matson, D. W. Schindler, and D. G. Tilman. 1997. Human Alterations of the Global Nitrogen Cycle: Sources and Consequences. *Ecological Applications* 7(3): 737–750.

Wagenhoff, A., C. R. Townsend, and C. D. Matthaei. 2012. Macroinvertebrate responses along broad stressor gradients of deposited fine sediment and dissolved nutrients: A stream mesocosm experiment. *Journal of Applied Ecology* 49(4): 892–902.

Wang, L., D. M. Robertson, and P. J. Garrison. 2007. Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: Implication to nutrient criteria development. *Environmental Management* 39(2): 194–212.

West Virginia Department of Environmental Protection (WV-DEP). 2015. Watershed assessment branch 2015 field sampling standard operating procedures. Division of Water and Waste Management, Watershed Assessment Branch, Charleston, WV. 496 pages. www.dep.wv.gov/WWE/watershed/wqmonitoring/Documents/SOP%20Doc/2015WABSOP/2015%20WAB%20Field%20Sampling%20SOP.pdf

Woznicki, Sean A., Nejadhashemia, A. Pouyan, Ross, Dennis M., Zhang, Zhen, Wang, Lizhu, and Esafahanian, Abdol-Hossein. 2015. Ecohydrological model parameter selection for stream health evaluation. *Science of the Total Environment* 511: 341-353.

Woznicki, Sean A., A. Pouyan Nejadhashemi, Mohammad Abouali, Matthew R. Herman, Elaheh Esfahanian, Yaseen A. Hamaamin, and Zhen Zhang. 2016. Ecohydrological Modeling for Large-Scale Environmental Impact Assessment. *Science of The Total Environment* 543(A): 274–86.

Yuan, L. L., and S. B. Norton. 2004. Assessing the relative severity of stressors at a watershed scale. *Environmental Monitoring and Assessment* 98: 323-349.

Zhou, Y., D. Scavia, and A. M. Michalak. 2014. Nutrient loading and meteorological conditions explain interannual variability of hypoxia in Chesapeake Bay. *Limnological Oceanography* 59(2): 373-384.